Design, Implementation & Analysis of a Low-Cost, Portable, Medical Measurement System through Computer Vision

by

Gareth van der Westhuizen

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Engineering (Mechatronic) at Stellenbosch University

Department of Mechanical and Mechatronic Engineering,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Supervisor: Prof. K. Schreve

March 2011

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: .....

Abstract

Design, Implementation & Analysis of a Low-Cost, Portable, Medical Measurement System through Computer Vision

G. van der Westhuizen

Department of Mechanical and Mechatronic Engineering,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Thesis: MScEng (Mechatronic)

March 2011

The In the Physiotherapy Division of the Faculty of Health Sciences on the Tygerberg Hospital Campus of the University of Stellenbosch, the challenge arose to develop a portable, affordable and yet accurate 3D measurement machine for the assessment of posture in school children in their classroom environment. Currently Division already uses a state-of-the-art VICON commercial medical measuring machine to measure human posture in 3D in their physiotherapy clinic, but the system is not portable and is too expensive to cart around to different places for testing.

To respond to this challenge, this Master's thesis designed and analyzed a machine and its supporting system through both research on stereo-vision methodologies and empirical appraisal in the field. In the development process, the research was required to overcome the limitations posed by small image resolutions and lens distortions that are typical of cheap cameras. The academic challenge lay in the development of an error prediction model through Jacobian derivation and Error Propagation Law, to predict uncertainties of angular measurement calculated by the system.

The research culminated in a system that is comparable in accuracy to the VICON within 3 mm, and that has 1.5 mm absolute accuracy within its own system

for a measurement volume radius of 2.5 m. As such, the developed error model is an exact predictor of the angular error to within 0.02° of arc. These results, for both system accuracy and the error model, exceed the expectations on the basis of the initial challenge of the system.

The development of the machine was successful in providing a prototype tool that is suitable for commercial development for use by physiotherapists in human posture measurement and assessment.

In its current incarnation, the machine will also serve the Engineering Faculty as the most fundamental form of a three-dimensional measuring apparatus using only basic theories and algorithms of stereo-vision, thereby providing a basic experimental platform from which further scientific research on the theory and application of computer vision can be conducted.

Uittreksel

Ontwerp, Toepassing & Analise
van 'n Goedkoop, Draagbare Mediese Meettoestel
deur Rekenaarvisie

("Design, Implementation & Analysis of a
Low-Cost, Portable, Medical Measurement System through Computer Vision")

G. van der Westhuizen

Departement Meganiese en Megatroniese Ingenieurswese,
Universiteit van Stellenbosch,
Privaatsak X1, Matieland 7602, Suid Afrika.

Tesis: MScIng (Megatronies)

Maart 2011

Die Fisioterapie Afdeling van die Fakulteit Gesondheidswetenskappe op die Tygerberg kampus van die Universiteit van Stellenbosch gebruik 'n allernuutste VICON kommersiële mediese meettoestel om menslike postuur in drie dimensies te meet.

Vanuit hierdie Afdeling het die uitdaging ontstaan om 'n draagbare, bekostigbare, maar tog akkurate, drie-dimensionele meetapparaat geskik vir die meet van die postuur van skoolkinders in die klaskamer te ontwikkel.

In aanvaarding van hierdie uitdaging, het hierdie Magistertesis 'n toestel en ondersteuningstels ontwerp en ontleed deur beide navorsing in stereo-visie metodiek en terplaatse beoordeling. In die ontwikkelingsproses moes die navorsing die beperkings wat deur klein-beeld resolusie en lens-distorsie (tipies van goedkoop kameras) meegebring word, oorkom. Die akademiese uitdaging lê in die ontwikkeling van 'n voorspellende foutmodel deur van die Jacobianse-afleiding en die

Fout Propageringswet gebruik te maak om onsekerheid van hoeksberekening deur die stelsel te voorspel.

Die navorsing het gelei tot 'n stelsel wat binne 3 mm vergelykbaar is in akkuraatheid met dié van die VICON en ook 1.5 mm absolute interne akkuraatheid het in 'n meet-volume radius van 2.5 m radius. Die ontwikkelde foutmodel is dus 'n presiese voorspeller van hoekfout tot binne 0.02° van boog. Die resultate met betrekking tot beide die akkuraatheid en die foutmodel het die oorspronklike verwagtinge van die uitdaging oortref.

Die ontwikkeling was suksesvol in die skep van 'n prototipe-toestel geskik vir kommersiële ontwikkeling, vir gebruik deur fisioterapeute in die meting en evaluering van menslike postuur.

Die stelsel is in sy fundamentele vorm, deur die gebruik van slegs basiese teorieë en algoritmes van stereo-visie, funksioneer as 'n drie-dimensionele meetapparaat. In die fundamentele vorm sal die stelsel die Ingenieursfakulteit dien as 'n basiese eksperimentele platform waarop verdere wetenskaplike navorsing in die teorie en toepassing van rekenaar-visie gedoen kan word.

Acknowledgements

To the following people and organizations:

- Prof. Kristiaan Schreve, for support with the project - adding insight and guidance wherever required - and for ensuring adequate funding for the project in the form of bursary and equipment funding.
- The National Research Foundation, for providing the funding for the project.
- The University of Stellenbosch, Faculty of Engineering and Department of Mechanical & Mechatronic (M&M) Engineering - even though it would seem self-evident, but the reality is that without this institution this work would not have been possible.
- Prof. Quinette Louw, Yolandi Brink and the Physiotherapy Division in the Faculty of Health Sciences of the University of Stellenbosch, for guidance in the development of the project where relevant, and for support with the experimentation.
- The technical support staff at the Dept. of M&M, for their expertise and ability to construct the machines and parts designed.
- Prof. B Herbst & Dr M Maritz of the Dept. of Applied Mathematics, for expertise and insights into Computer Vision & Image Processing.
- Carlo Schenke, foremost of my friends, who lent direct physical support during the experimentation, as well as being a second brain to bounce ideas off.
- Dr DNJ Els, for constant advice and updates for LaTeX, and for knowing the latest styles in formatting and how to draw crisp graphics.
- My parents, beyond the financial support, love & understanding, simply for being there for me – for believing in me all along.

- To my close friends & family; from lending a hand to glue markers, to being guinea-pigs and proof-readers, to simply bringing me that next cup of tea...

my sincere thanks for all the help you all have given at various times through-out.

This thesis was typeset in LaTeX using the usthesis package of DNJ Els, as well as various other LaTeX packages. Notably, the references were compiled in BibTeX using the usbib package of Els. The TeX was typed in WinEdt TeX editor.

All images are that of the author. Photo-editing was accomplished in GIMP, graphics were created in either Inkscape or OpenOfficeDraw and graphs were generated in Excel and Matlab using encapsulated post-script (EPS) formats.

To the elusive thought; the one that "$\alpha\ \wr^{\omega}\ \mathring{a}_{y}\ \int$" escapes me...

Contents

List of Figures

List of Tables

Nomenclature

Vectors & Matrices

$\bar{X}$	The arithmetic mean of X
$\lvert A \rvert$	The determinant of matrix A
$\cdot$	The dot-product, matrix multiplication, $C = A \cdot B = AB$
$\div$	Element-wise division, $C = A \div B : C_{ij} = A_{ij}/B_{ij}$
$\bullet$	Element-wise multiplication, $C = A \bullet B : C_{ij} = A_{ij} \cdot B_{ij}$
$\lVert A \rVert$	The L_2 norm of A
$\boldsymbol{A}$ or A	Matrix named A
$\emptyset$	The null-matrix or null-vector
$\langle x_1, \cdots, x_n \rangle$	Row-vector of n-elements
$\vec{X}$ & $\vec{x}$	Vectors named X & x

Scalars

μ_x or $\bar{x}$	The arithmetic mean of x
x	Scalar named x
σ_x	The standard deviation of x
σ_x^2	The variance of x

Units

fps	Frames per second . . . . . [frames/s]
px	Pixel - smallest portion of a raster image . . . . . [px]

Subscripts

$_{x,y,z}$	Axes corresponding to the unit vectors $\left[\hat{i}, \hat{j}, \hat{k}\right]$

Superscripts

${}^{E}\vec{A}$ Coordinate reference frame, vector $\vec{A}$ is in reference frame E

A^{T} Transpose of A.

Phrases, Definitions & Contractions

CCS Camera coordinate system – the three dimensional reference frame of a camera relative to the camera body.

CMM Computer Measurement Machine – one of the many machines that accurately measure objects in the Metrology Lab

ICS Image coordinate system – the two dimensional reference frame relative to the image plane of the camera.

GCS Global coordinate system – the three dimensional reference frame which encompasses the measurement volume as well as the cameras and the researchers.

PMM Posture Measurement Machine – the design objective of this thesis

VICON The high-level commercial system for biomedical motion analysis via multiple camera computer vision

2D Two dimensional

3D Three dimensional

$\mathbb{P}^n$ Projective space of dimension n

$\mathbb{R}^n$ Real space of dimension n

Chapter 1

Introduction

1.1 Background to the Research

Currently, the National Research Foundation (NRF) is funding research by the Physiotherapy Department of Stellenbosch University to study the causes of postural discomfort in school-going adolescents.

A critical aspect of the physiotherapy research is posture measurement. Obviously, the more accurately the posture can be measured, the more accurately the patient can be modelled.

Currently, a well established and very versatile method of non-invasive, non-contact posture measurement is the use of imaging systems that use specifically placed markers on the patient. Images of the patient are then recorded from cameras and either a two-dimensional analysis or a full, three-dimensional reconstruction is performed on the captured data. From this the joint and segment and angles are calculated; these describe the posture of the subject. The advantages of this technique are that the data is a precise snapshot in time, and the kinematics of the patient in that instant would be very small (especially compared to a person measuring the patient with a ruler). Further, because it is an instant in time, the patient would not have to artificially hold a specific pose. Disadvantages of the technique are that the currently available commercial imaging machines are extremely costly (in excess of R100 000) and cumbersome.

Ideally, the measurements of the patient should be taken within the environment in which the patient is most active. To this end, the imagining system should be easy to move to the place of study (e.g. schools at which the students spend the most time), as well as be easily configurable so as to be placed in the test area with

minimal effort or rearrangement.

In her Master's thesis, the physiotherapist Yolandi Brink, carried out posture research using planar angle definitions that were calculated from markers picked in two-dimensional (2D) images of the subjects [1]. In doing so, Brink worked with Prof. K. Schreve of the Mechanical and Mechatronic Engineering Department of Stellenbosch University.

It was in this collaboration that the niche for the design of a low-cost, portable, three-dimensional posture measurement tool was identified. Applications and proposals for funding were submitted to the NRF, and this kicked off this Master's degree research which would tie in with the physiotherapy research testing learners' posture at schools.

This Master's thesis is based on the development and implementation of a cost-effective and portable three-dimensional, multiple-camera measuring system. Emphasis is not placed on the development, but rather on the implementation, thus the use of readily available, off-the-shelf products or currently available equipment that can be found in the Department. Further, because this is a research-based Master's thesis, there will be further research specifications and requirements that will tie in with the machine so that it may be used to further explore the field of computer vision while collecting relevant data for the physiotherapists.

Recently, the Division of Physiotherapy took ownership of a VICON MX Motion Analysis System, which uses cameras to track markers on a subject and then reconstruct a fitted 3D model. The technology herein is very advanced and very expensive (in the order of millions of Rand); as a result, it is not an option to take the VICON system along to the schools. However, it is available for use as a comparative system against which to test the system being designed.

1.2 Problem Statement

The aim of this Master's thesis was to design and implement a computer vision-based posture measuring system that is portable, accurate and inexpensive. In doing so, the aim was to determine how improvements affect the system and to find an analytical model for the expected angular error.

A breakdown of this problem statement follows, from its origins in the client specification, through its evolution into the engineering specification, and showing scope, limitations and boundaries.

1.2.1 Physiotherapy "Client" Specifications

From discussions with the physiotherapists it was found that they were very relaxed in their specifications, only requiring that the following be fulfilled:

- Must not encroach upon the subject.
- Must be accurate to within $\pm 5°$ for angular measurements.
- Must be repeatable.
- Must be portable.
- Must be useable by people with limited familiarity with computer equipment.
- Need only capture a single instance – not full motion sequences.

In particular, Yolandi Brink is lead Physiotherapist for whom the PMM as physiotherapy tool is being designed. Brink not only contributed the client specifications but also assisted with some of the design iterations and accuracy testing of the system.

1.2.2 Material & Research Constraints/Requirements

Departmental constraints were focussed largely on which materials were to be used. In order to keep the costs down, it was preferable to use the currently available Point-Grey research cameras. Ideally this system would use simple, consumer web-cams; the use of web-cams as opposed to specialist cameras would greatly help reduce the cost of the system. While these research cameras are more advanced than conventional web-cams, they are not so advanced that they are incomparable to web-cams. Further, the flexibility afforded by the software development kit (SDK) and the fact that, as research cameras their internal workings are accessible to the design engineer, favoured the use of these cameras.

Camera tripods were already available; being the cheap, generic type, these were favoured to see if they were suitable given their low cost. Cross-bars, clamps and so forth had to be made (where possible) from scrap metal available in the SMD workshop. Redesign and customization had to be kept to a minimum so as to make the system as *off-the-shelf* as possible.

A software constraint was that the system had to be implemented in Open-Source Python. This would enable it to use existing modules written by the supervisor and allow it to be integrated in future projects of the computer vision group, which uses Python as its standard platform.

1.3 Research Aims

Computer vision is no different from any research field in that the completion of current research leads to the need for more research. While computer vision is not exactly a young field, being written about in the 1970s already, it is the advent of increasingly powerful modern desktop computing systems that has opened the field of computer vision to broader research. As such, there is no shortage of topics that need to be investigated.

1.3.1 Within the Design Process

The idea is to design the system as *loose* as possible and then to systematically investigate what further improvements increase the accuracy of the system. Accordingly, focus was first applied to the elements that gave rise to the greatest increase in accuracy.

1.3.2 Error Modelling

Once improvements have been made and the system is running, there will be methods that measure the accuracy of the system – after all, if the system cannot be verified as accurate then it really does not serve a purpose in the first place.

But, beyond that, there should be some way in which the accuracy of the system can be predicted – especially in a way that is relevant to the end-user – in this case the physiotherapists.

Thus, it would be useful to be able to express the system's accuracy in terms of angular measurements between points. In this way the applicability of the system to a specific type of study can be predicted even before measurements have commenced.

1.4 Scope

The time frame of this Master's thesis was set for eighteen months, from May 2009. The time was allocated so that during the design of the machine, the machine would be minimally up and running within six to seven months; that is the cameras should be able communicate and stream data with the computer, rudimentary marker selection should be available to allow calibration and unrefined reconstruction for initial testing purposes. A further four to five months allocated for programming the software to retrieve data. Thereafter there would be three months for improvements to the machine and research on the error model. As a result, the project would run for thirteen to fifteen months (leaving a month's grace time), with two months to collate the data and prepare the thesis.

1.4.1 Design

The only upper limit on the design scope – apart from keeping the technology affordable – was that each improvement had to be almost fully implemented and confirmed to make an improvement to the system before the next improvement on the system could be started. This meant that, if there was time available, as many improvements as possible could be attempted on the design. By the end of the Master's thesis, any incomplete improvements would not be implemented, but reserved as suggestions for further prototyping of the system.

1.4.2 Research

Foremost, the research seeks to demonstrate whether each suggestion was an actual improvement upon the system. Thus, these improvements will need to be quantified and compared to each other and to the most basic implementation of the system.

The error model predicts the errors in angular measurements from the machine. It does not investigate how errors propagate through the camera system itself, but rather uses such predications as inputs for estimating the angular error. The predication of errors through the camera system will be estimated using existing methods in computer vision.

Again, the initial limitations on the scope of the research were very open ended, with the plan being to attempt the initial considerations and, if these are completed

in time, to consider taking them further as applicable within the time constraints.

1.5 Thesis Overview

From here the thesis will run through the literature review to highlight both applicable work and works of interest that have been done.

Thereafter, §2 will deal with the machine and supporting system – how it is physically constructed and what software it runs – as well as a look into the underlying mathematics that it runs on. Subsequently, improvements made to the system to bring it inline with the required specifications are discussed in §3. The research aspects of the error model are presented in §4. While §5 highlights how the system was used by the physiotherapists as part of their study in schools. Most importantly, the accuracy of the system is investigated in §6, through an observational experiment and analysis of the findings.

Finally, the whole thesis is concluded with §7 which looks back at the thesis in its entirety; presenting a summary of achievements, mentioning short-comings and remarking on a few points that are worthy of further, future, consideration.

1.6 Literature Review

In his 1997 book, *How the Mind Works*, experimental psychologist Steven Pinker took a deep look at the human mind and its supporting systems. Of note for this thesis is a look at how vision works in humans: from data acquisition by the eye to signal and image processing in the brain in order for the world to be visually understood by the human. This is a process that is taken for granted by living creatures as they go about their daily lives.

The human mind is capable of vast jumps in logic and of parallel processing, enabling humans – in fact anything with a brain and eyes – to navigate their surroundings using visual information and optical cues. However, trying to replicate this in computer systems is an entirely different problem, one which has created the field of computer vision.

Inherently, building a machine does not necessarily require the creation of new techniques for completing a task when adequate methods exist. In fact, in this case, the machine is to serve as the most basic implementation of a stereo-vision-based measuring machine, one that can achieve the specifications set by the physiother-

apists. In doing so, the machine will serve as the base level of a research platform, from which further developments or more advanced techniques can be compared to the simplest – yet accurate – implementation of such a system. While this thesis relies heavily on the theory and algorithms presented by Hartley and Zisserman [2] in *Multiple View Geometry* and by Faugeras [3] in *Three-Dimensional Computer Vision: A Geometric Viewpoint*, it would be naïve not to familiarize oneself with the current research in the field of computer vision and, more specifically, stereo-vision.

1.6.1 Stereo-Vision System

Stereo reconstruction of 3D data is not limited to single-image, stationary cameras. In 1982, Whittle [4] showed that accurate 3D reconstruction could be achieved for kinematic analysis using calibrated, stationary television cameras. Subsequently, in 1989, Weng *et al.* [5] presented an algorithm to derive motion and structure from two perspective views. Of interest is that they showed that depth away from the camera, along the optical axis, affects accuracy more significantly than motion in the camera plane, perpendicular to the optical axis.

More recently, in 2010, Yang *et al.* [6] presented an error model that uses the strengths of both motion- and stereo-reconstruction to better correct for errors in the final reconstruction. For the physiotherapy study the patient is required to be stationary; to make the system portable the camera system cannot employ a robust, intricate tracking system and thus must also remain stationary. The result is that the system cannot reap the benefits of motion reconstruction, but the trade-off is that the system is not hindered by the additional equipment.

1.6.2 Calibration

Without a proper calibration for the system, the camera matrices would be meaningless and the reconstructed values completely worthless. Understandably, calibration is one of the most researched aspects of stereo-vision. In keeping with the base-level concept for the machine, the machine will use a static ground-truth calibration object and the Direct Linear Transformation (DLT) algorithm, the precision of which is described by Chapron *et al.* [7]. Certainly the DLT is not the only method for calibration, it will be implemented in the system because it is a well documented method, it is a more basic method – meaning that it is easier to implement and use, and that this ties in with implementing the system as the base-level

of a research platform. Additionally, the DLT algorithm is already in use within the Department's Computer Vision Group through the use of pyMultiCam and to this end its usage is well known within the group. Other algorithms and methods for calibration that are available in the literature are as follows:

Pedersini *et al.* [8] describe a method for the accurate and simple calibration of a multi-camera stereo system. However, the calibration object is a 2D target set at difference angles, which limits the capture volume as it cannot be viewed from 360°.

In terms of calibration precision, Song *et al.* [9] defined a method for the calibration of cameras using the cameras in pairs. By projecting the cameras' world coordinates onto the image plane, they were able to more accurately calculate the rotations and translations of the cameras from the world origin. Then, by minimizing the reconstruction error of a calibration object, the distortion parameters of the cameras could accurately be derived, allowing for more accurate distortion correction. The subsequent reconstructions by this method were found to be accurate to within 0.01 mm, using cameras of 1280×1024 px (roughly twice the resolution of the cameras used in this thesis).

Of direct significance to this project is that Zhou *et al.* [10] found that the calibration error of the cameras' intrinsic parameters can be reduced when the calibration volume is designed to fill as much of the available image space in the cameras as possible. An order of magnitude improvement was found when using a full-view as opposed to a quarter-view. Both cases were calibrated using the DLT algorithm, making this an important consideration for planning the positions of the cameras in the physiotherapy study.

In the 1996, Zhou *et al.* [11] showed that an increased signal-to-noise ratio (SNR) in the image leads to an improved calibration. This allows for a basis from which to evaluate how accurate a given camera will be, based on its SNR.

Another method is that of *auto-calibration*, whereby the motion of discernable points in the scene is used to determine the parameters of the cameras. As described by Hartley and Zisserman [2] and Hartley [12], a point in the scene is captured, after which motion occurs (either by moving the cameras or the point in the scene), and then the point is captured again. The transformation between these two points allows for calibration of the camera parameters up to an affinity. The subsequent metric calibration requires a calibration object, or known geometry of features in the scene. The important difference is that the calibration object for

auto-calibration can be much smaller than the desired capture volume. It is the motion of the point in the scene that will fill the capture volume, defining it up to an affinity; the metric calibration can then be small due to the fact that Euclidian-Cartesian coordinates are linear by definition.

In 2002, Shashua and Navab [13] introduced the concept of the relative affine structure, that is an affine framework that unifies Euclidian, projective and affine frameworks to a viewer-centred invariant in a more simplified manner than previously found. In doing so, the time taken for the reconstruction of 3D from stereo-image streams is decreased. However, this method has not been tested for metric reconstruction.

Wu *et al.* [14] showed that it was possible to perform auto-calibration with a one-dimensional moving object. It was found that the 1D object in motion is in essence equivalent to a 2D object in planar motion [14].

A drawback of the project is that no easily accessible methods of auto-calibration were found to be implemented for the desired Python software platform. For the purpose of the physiotherapy study, auto-calibration would definitely increase the portability of the system.

1.6.3 System Uncertainty

Many of the papers already mentioned deal with the system uncertainty by showing how each method is an improvement; these papers should be taken into consideration when designing to minimize the overall uncertainty of the system.

The portion of the work by Zhou *et al.* [10] that improved camera parameter estimation through motion and stereo calibration was part of creating an error model for scene reconstruction from such methods. As such, their model analytically describes the statistical error expected on a marker position due to the intrinsic and extrinsic parameters of the cameras – both for motion and stereo as separate entities, and then as a combined technique. However, the uncertainty is determined for a stereo-pair with parallel focal-axes – thus it does not account for relative camera rotation in the pair.

In 1995, Trobina [15] showed that measurements parallel to the optical axis of a camera system had significantly larger errors than measurements perpendicular to the optical axis. In 2004, Cheong and Peh [16] derived a method to predict the amount of depth distortion in scene reconstruction as a result of calibration

uncertainties. This would enable a designer to selectively choose reconstructions of a point from multiple cameras based on preferring camera-pairs closer to that point and camera-pairs that are not viewing the point parallel to the optical axis.

Weng *et al.* [17] investigated the effects of camera distortion on reconstruction accuracy and how camera calibration affects the correction of distortion, and evaluated how the distortion affects the system and what sort of distortion model – radial, tangential, decentring, etc. – has the greater effect on lens systems.

Mayoral *et al.* [18] evaluated the effect of correspondence errors on stereo-reconstruction, an important source of uncertainty in reconstruction. Their work shows how the scene can be evaluated and then used to select the most appropriate method to calculate the correspondence error. The more accurately this error can be accounted for, the more accurate the system.

Chowdhury and Chellappa (2008) [19] investigated the statistical error propagation in 3D reconstruction, as it pertains to the context of video image streams and camera motion. They derived an error model for the bias and covariance of the entire system based on inputs to the camera system, intrinsic and extrinsic parameters, image error and camera motion. Though not directly applicable to this thesis, the methodologies followed bear investigation. Their model was analytically derived and showed that the 3D reconstruction is statistically biased, with the analytically expression showing that the bias is numerically significant. Thus, no matter how accurate the calibration and lens distortion correction, there will need to be an error correction for the reconstruction algorithms themselves. These findings correspond with those of Haralick [20], who in 1994 investigated how errors propagate through camera systems without deriving an analytical result.

Error models are important, not only because they can reduce complex structures into manageable concepts and can be used to correct errors post-fact. In terms of experimental design, the error model can be used to predict the expected error before the experiment has been carried out. Later in the thesis, an error model is derived for predicting the uncertainty associated with angles that are calculated from points reconstructed from 3D stereo-vision. While there is much research that has been done on the errors of the 3D reconstructed points, not much exists on what the effects of those errors are on subsequent calculations. The research desires to bridge this gap and, in doing so, will allow experimenters to better utilize stereo-vision systems when designing experiments to calculate angles (an important aspect in physiotherapy).

1.6.4 Reconstruction

Seitz *et al.* [21] present a comparison and evaluation of multi-view stereo-reconstruction algorithms. Even though they dealt more with the reconstruction of a mesh-surface so that un-markered objects could be reconstructed successfully in 3D, they also dealt with the multiple camera system. In their system they used 47 cameras to form a viewing hemisphere around the object. They concentrated on point correspondences (photo-consistency measures), visibility models and reconstruction algorithms. They also found that there were varying offset errors between algorithms, as proven by Chowdhury and Chellappa [19].

Fusiello *et al.* [22] show how compactly the rectification of stereo-pairs can be achieved. More than anything, their paper serves to show how succinctly these algorithms can be implemented and how successfully the field of stereo-vision has been moving forward.

Yuan and Ryd [23] presented a computer simulation of 3D marker reconstruction in Röntgen Stereophotogrammetric Analysis (RSA) – an x-ray imaging system. Even though the x-ray system is different, the underlying mathematics is the same, and Yuan and Ryd present an interesting look at how such a system could be simulated.

In the end, as far as reconstruction is concerned, it is Olague and Mohr (2002) [24] who have the most significant direct impact on this study. Their work dealt with the optimal placement of cameras for accurate reconstruction. This allows the experimenters to design their capture volume for the best available accuracy or to choose a different area as a capture volume to better utilize the cameras. The work of Olague and Mohr ties in well with that of Zhou *et al.* [10], as discussed earlier, concerning the need to fill the camera view with as much of the capture volume as possible.

The use of the trifocal tensor unfortunately falls completely beyond the mathematical scope of this thesis. However, as discussed by Hartley and Zisserman [2], it is a powerful method that should not be overlooked in multi-camera systems.

1.6.5 Conclusion

As can be seen there is no shortage of relevant information available on computer vision. In all these papers are to augment the thinking and ideas of the thesis. The purpose of the thesis is not to replicate any of these papers.

Chapter 2

The Posture Measurement Machine (PMM)

The system to be designed for the physiotherapists has two distinct parts that interact to form the whole: the *physical system* and the *software*. This chapter will discuss these parts and their sub-systems, as well as the operating procedures of the machine. The machine will be referred to as the Posture Measurement Machine or PMM throughout this thesis.

2.1 Physical System

The physical system consists of different sub-systems and parts that are used to form the overall Posture Measurement Machine (PMM).

2.1.1 Parts List

The following parts form the physical system:

Cameras: Five 0.3 MP CMOS FireFly MV – 640 × 480 – Point Grey Research (IEEE 1394a "Firewire").

Lenses: 6 mm fixed focal length – Point Grey Research.

IEEE Hub: 5-port and 3-port hubs, connect multiple cameras to a single IEEE bus.

EXP Card: IEEE 1394b Firewire bus expansion card for laptop.

Power Supply Supplies required voltages to expansion card (thus to hubs and cameras).

Computer: The current system runs on the Windows operating system. Configuration confirmed to work in Windows Vista.

Cross-bars: Two cross-bars onto which cameras are mounted.

Tripods: Two tripods that support the cross-bars.

Clamps: Two clamps that connect the cross-bars to the tripods.

Figure 2.1 shows the basic parts of the physical system:

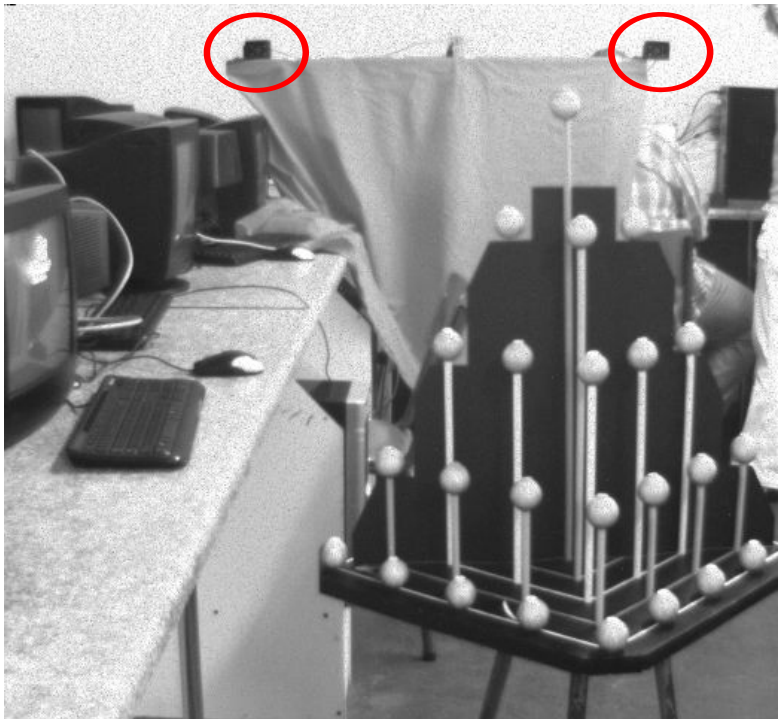


Figure 2.1: Photograph of the measurement machine showing the highlighted cameras, the computer to the left and the calibration object to the right.

2.1.2 Configuration

The system is connected as shown in Figure 2.2:

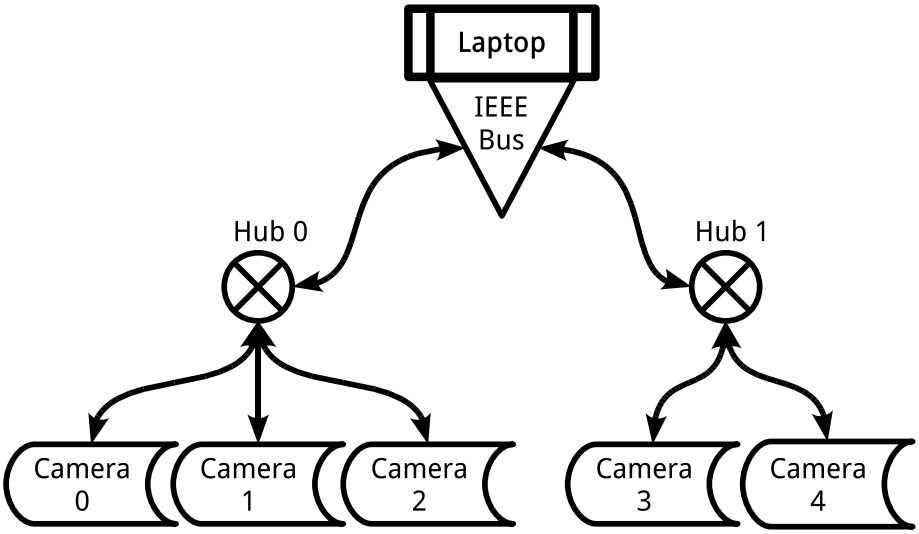


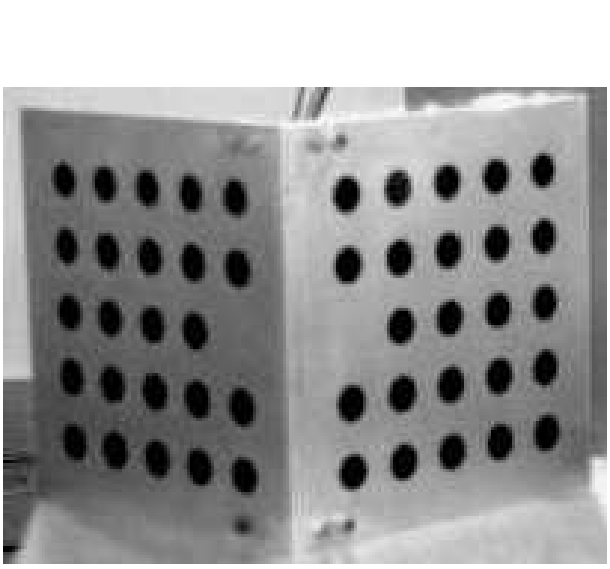
Figure 2.2: System connections

The camera housings are mounted onto the cross-bar by L-brackets and bolts. This allows the cameras to be individually rotated horizontally by loosening the bolts and manually turning them. The cross-bar is mounted on the tripods by clamping them to the tripod head. The bar as a whole can be rotated horizontally and tilted vertically by manipulating the tripod head.

The first critical piece of equipment, a camera outside of its casing, is shown in Figure 2.3. The lens is approximately 12.5 mm in diameter. As can be seen, these are not large nor bulky cameras.

Figure 2.3: Firefly MV camera

The second most important piece of equipment is the calibration object. The two different objects used are shown in Figure 2.4a (as seen by the Firefly camera) and Figure 2.4b.



(a) Aluminium corner

(b) Pyramid sticks

Figure 2.4: Calibration ground-truth objects form the basis of system calibration. The markers on these objects must be known independently of the system and to significant accuracy in order to have an accurate system.

2.2 Software

The software system of the PMM was implemented in Python[1] (version 2.6), using the OpenCV[2] computer vision bindings, PyOpenGL[3] graphics library bindings, VPython[4] module and pyMultiCam [25] module.

The interface is command-line driven, with the user linearly supplying input data as requested by the system. Input data, such as images, world points and marker definitions, is stored in specific, separate folders. Images are stored in JPEG format, while all other data is stored as comma-separated values in text files. Data output is as comma-separated values in text files into an output folder.

User-supplied data input includes the number of cameras active in the capture and identification names for the capture calibration, and for the individual subjects subsequently captured using the given calibration.

Additional software that is used is the Point Grey FirePro software development kit (SDK). This not only incorporates the hardware drivers to interface the operating system with the cameras, but a program specifically written to synchronize all the cameras present across multiple IEEE busses that are present on the computer.

[1] http://www.python.org
[2] http://opencv.willowgarage.com/
[3] http://pyopengl.sourceforge.net/
[4] http://vpython.org/

2.3 Overview – How the System Works

The operation of the machine can be separated into distinct stages, as illustrated in Figure 2.5.

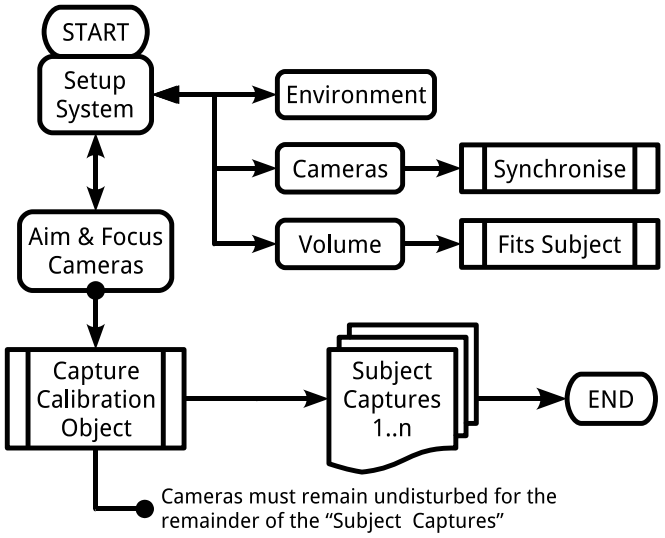


Figure 2.5: Overall flow for capturing a subject

The camera system of the PMM is set up, then carefully calibrated for the viewing *volume*; the calibration produces a camera matrix per camera. A markered subject[5] is placed inside this calibrated volume and synchronized images are captured from different viewpoints around the volume. For each image, the image coordinate value for each marker visible in the image is found and recorded. The cameras are then paired and each pair's image coordinates and camera matrix are used to triangulate the global coordinates of each marker visible by the pair. Once all the markers are triangulated per pair, the total set of marker values is combined to produce a single set of global coordinates for all the markers on the subject. From these values the posture angles and lengths can be calculated as required. A piece-for-piece breakdown of the operating procedures of the PMM is given in §2.5.

2.4 A Look at the Fundamentals of Stereo-vision

The theoretical base on which the PMM operates is *Stereo-vision*, which describes how real world points in 3D space project onto 2D image planes and, if the orient-

[5] A subject with markers in place.

ations and positions of these 2D planes are known, how points on these planes can be used to reconstruct the original 3D point.

For all systems in the real, spatial world there is always some form of coordinate system. The PMM has an origin at a point $O \in \mathbb{R}^3$. The cameras have a coordinate system relative to their own rigid body. The origin of the camera in the global coordinate system (GCS) is given by $\vec{C} \in \mathbb{R}^3$, called the *camera centre*. Further, the axes of the camera coordinate system (CCS) need not be aligned with the GCS – this stands to reason, as a camera may be arbitrarily rotated in the GCS. On the camera itself, an axis is usually aligned along the centre of the lens. This is the *focal* or *principal* axis and is perpendicular to the image plane of the camera. Any point in $\mathbb{R}^3$ can be represented on the image plane by projecting it through the camera centre. This can be seen in Figure 2.6, where the point $\vec{X}_1 \in \mathbb{R}^3$ is projected as $\vec{x}_1 \in \mathbb{P}^3$.

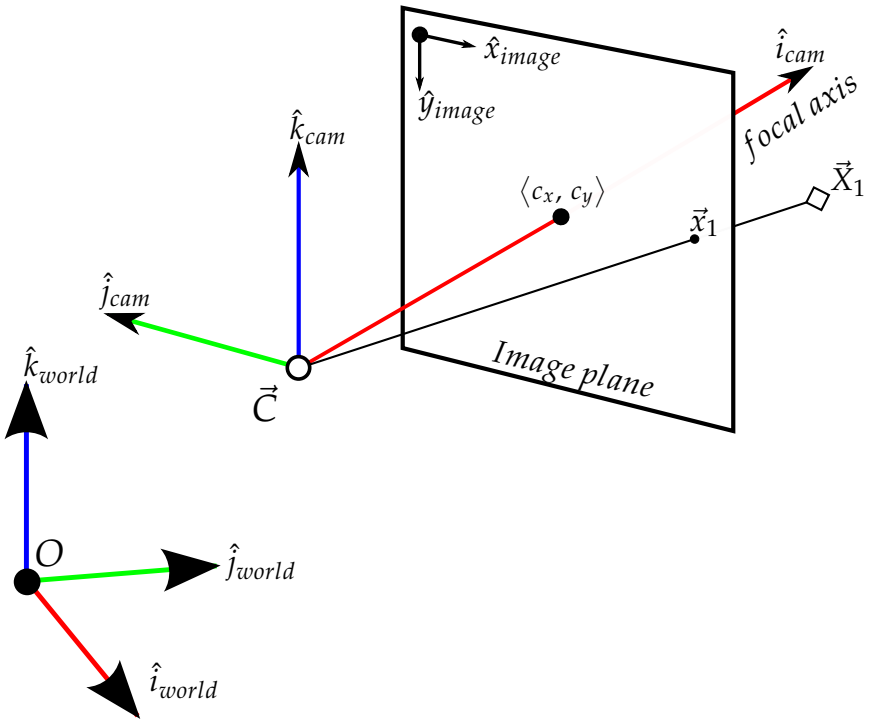


Figure 2.6: The projective spatial orientations of a camera, showing the global coordinate system ($\boldsymbol{O}$), the camera centre ($\vec{C}$), the camera coordinate system, the image plane, the image coordinate system and the the projection of a real point ($\vec{X}_1$) onto the image plane ($\vec{x}_1$).

Now, there exits a homography, $\boldsymbol{P}$, that maps the real world point, $\vec{X} \in \mathbb{P}^4$, to

the image point, $\vec{x} \in \mathbb{P}^3$ – therefore $\boldsymbol{P}$ must be a 3×4 matrix.

$$\vec{x} = P\vec{X} \tag{2.1}$$

The homography P is called the *camera matrix* and encompasses a number of things: a calibration matrix, $\boldsymbol{K}$, a rotation matrix, $\boldsymbol{R}$ and the camera centre vector, $\vec{C}$, such that:

$$\boldsymbol{P} = \boldsymbol{K}\boldsymbol{R}[\boldsymbol{I}|-\vec{C}] \tag{2.2}$$

The rotation matrix defines how the camera coordinate system is rotated in the global coordinate system, the camera centre defines where the camera is positioned in the real world, and together they form the *extrinsic* parameters of the camera.

The calibration matrix, $\boldsymbol{K}$, contains all the *intrinsic* parameters of the camera: the focal length, f, the projection of camera centre onto the camera plane[6], $\langle c_x, c_y \rangle$, and the metrication parameters, m_x and m_y. The metrication parameters scale how the pixel distances translate into global coordinate units. As the calibration object is defined in millimetres the resulting measurements will be in millimetres.

$$\boldsymbol{K} = \begin{bmatrix} m_x & 0 & 0 \\ 0 & m_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2.3}$$

Determining the camera matrix by finding each of the intrinsic and extrinsic parameters is possible, but extremely difficult. $\boldsymbol{P}$ is sensitive to small changes in each parameter, and trying to determine each parameter may often not even be feasible, because *any* change in the position of the camera will change several of the parameters. Focal length, rotation and position will all need to be remeasured should the cameras be moved. Further, physically measuring each of these elements accurately for multiple cameras would take vast amounts of time, etc. Fortunately there is an easier way, namely calibration, as discussed in §2.5.2. This shows how, given a special set of world points ($\vec{X}_n$) and corresponding image points ($\vec{x}_n$), the camera matrix can be solved.

An important point about stereo-vision is that it requires more than one camera — a real world point cannot be fully reconstructed from one image plane. Therefore there must be at least one more camera calibrated within the same world system, using the same calibration object. Cameras calibrated in this way would then

[6]The projection of the camera centre onto the camera plane is determined as part of the physical design of the camera; where possible it is available from the manufacturer.

form pairs with each other, and these pairs would be used to triangulated the real world point, as discussed in §2.5.6.

2.5 Operating Procedures

2.5.1 Setup

Depending on what needs to be captured, the volume is cleared of unnecessary objects; the remaining objects are positioned as required. The cameras are set up on the tripods and positioned so that they see as much of the volume as possible, and so that all markers are visible by at least one pair of cameras.

Once the cameras are connected to the computer, they must be focused and the viewing volume can be refined. It is important that the cameras are now synchronized on the system.

At the same time the subject will have special reflective markers placed on the points that need to be measured. The markers are covered in 3M Scotchlite, a retro-reflective tape. Any light that hits the tape will be reflected back in the opposite direction (as close to) as what it initially arrived along. It is important to double check that all markers on the subject can be seen by at least two cameras and that the calibration object can be seen by all the cameras.

Figure 2.7 shows a top view of a typical camera arrangement around a subject. The cameras are basically positioned in a hexagon to ensure that all markers are captured from the different viewpoints. Certain markers need only be seen by certain camera pairs, and it is important that all markers are seen by at least one camera pair.

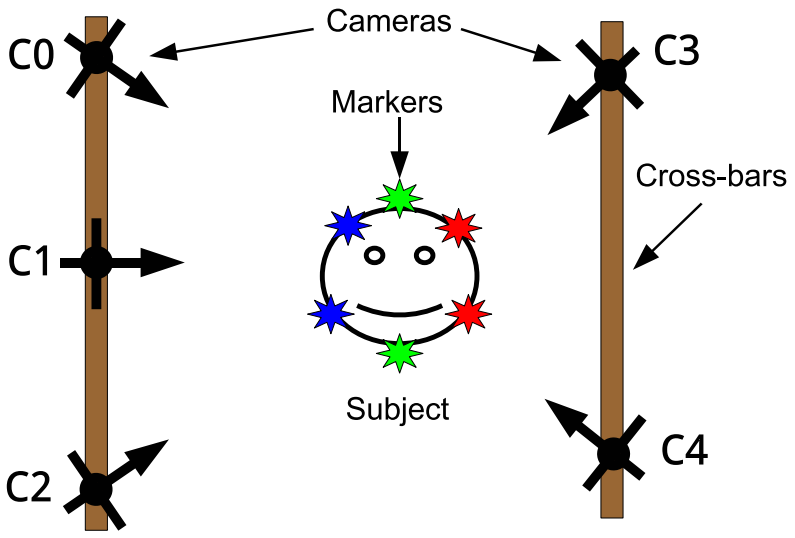


Figure 2.7: Camera arrangements in the top view around the subject.

2.5.2 Calibration

Calibration is performed with a calibration object. The object is specially built for this purpose in that the positions in the GCS (world points $WP \in \mathbb{R}^3$) of specific landmarks on the object are known to sub-millimetre accuracy; as measured on a computer measurement machine (CMM). These values are stored in a file accessible by the program and are used by the calibration algorithm.

The calibration object is placed inside the volume and captured to obtain images of the calibration object in the volume. From these images, the image coordinates (image points $IP \in \mathbb{P}^3$) of the landmarks can be found.

The Gold Standard calibration algorithm [2; 25] was implemented. The algorithm determines the camera matrix ($\boldsymbol{P}$) by using a direct linear transform (DLT) to create an estimated camera matrix, $\widetilde{\boldsymbol{P}}$, and then reduces the geometric error, ε, between the image points and the transformed world points, $\varepsilon = \left\| IP - \widetilde{\boldsymbol{P}} \cdot WP \right\|$. The $\widetilde{\boldsymbol{P}}$ that corresponds to the smallest possible geometric error is taken as the camera matrix, $\boldsymbol{P}$.

The implementation in pyMultiCam can also return the linear distortion parameters $\{\kappa_1, \kappa_2, x_c, y_c\}$, discussed in § 3.2. Figure 2.8 shows the generic process for calibration as implemented in the system.

Further, the coordinate system of the calibration object will become the coordinate system of all markers triangulated using that specific calibration – this includes axis, origin, orientation and unit scale. This means that choosing the right

calibration object for the desired capture is also an important consideration. Improvements and changes to the calibration object, as well as reasons for the current method, are discussed in §3.1.

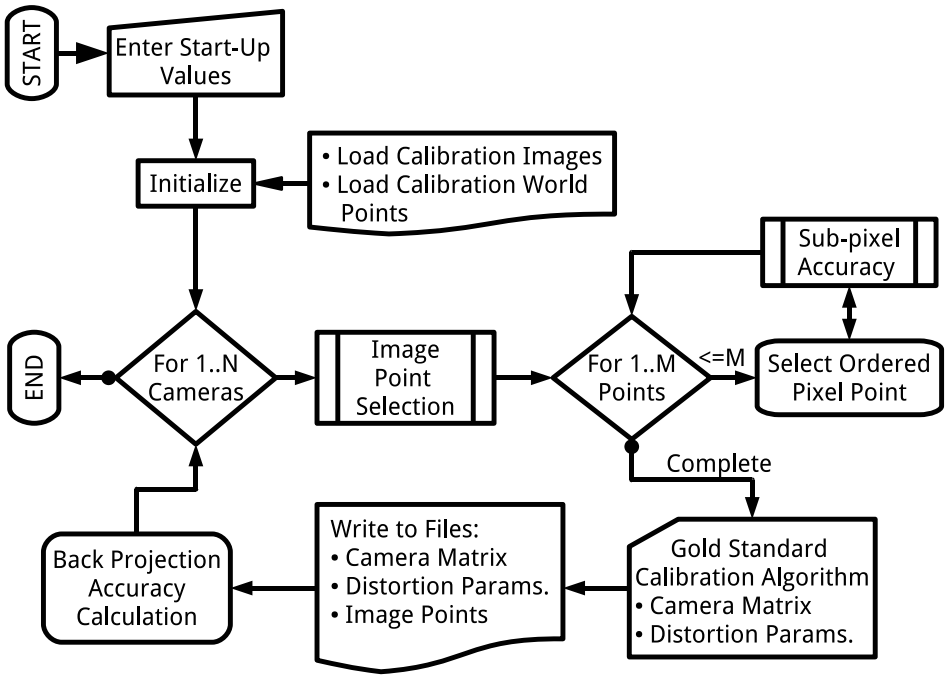


Figure 2.8: Program flow for calibration

2.5.3 Marker Placement

The marker placement is not only limited to the needs of the physiotherapists, for a vision based measuring system the markers can be placed at any point that is to be measured. The definition of where the markers are to be placed on the capture subject forms the *capture model*. However, as this is a medical measurement system, the model is selected by the physiotherapist from established models and the resulting marker placement corresponds to very specific biological landmarks on the subject. Marker names that correspond to marker selection order are defined by a text file. The marker model (as selected by the physiotherapist) should be well defined and remain constant and consistent throughout the study.

In the case of this project, the model was defined by the physiotherapist, corresponding to Figure 2.9:

#	Name	Description
0	LCanth	Left Canthus
1	RCanth	Right Canthus
2	LTrach	Left Tragus
3	RTrach	Right Tragus
4	C7	7^{th} Cervical Vertebra
5	T5	5^{th} Thoracic Vertebra
6	Strn	Sternum
7	LHip	Left Greater Trochanter
8	RHip	Right Greater Trochanter

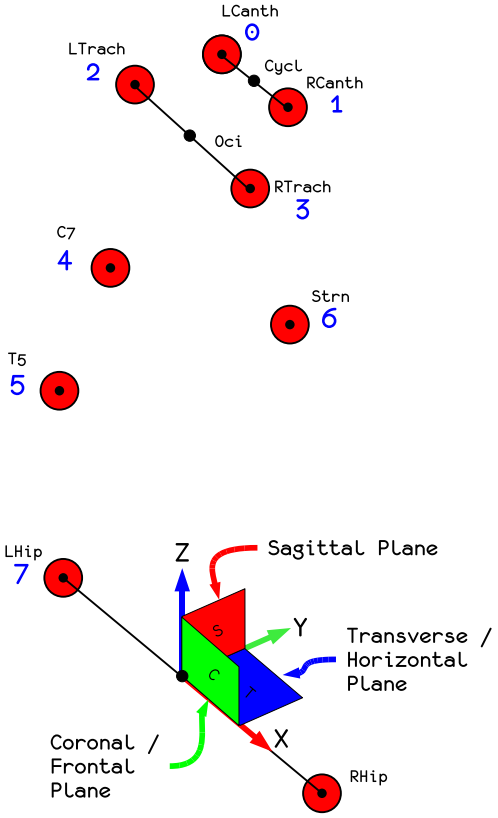


Figure 2.9: Model definition: Marker positions and numbers

Extra control markers can be added as needed in case the triangulation requires more markers than are available on the model; these extra markers do not need to be added as formal definitions in the model.

2.5.4 Subject Capture

The capturing of the subject is the same procedure for both calibration image capture and actual subject image capture. The software that currently performs this task is a demo program shipped with the Point Grey SDK ("`MultipleCameraWriteToDiskEx`"). Currently the use of the demo program is to simplify the design process as the program, ensures that all the images captured in a short time interval (approximately 15 s) have been captured synchronously.

The program runs long enough to ensure that 100 synchronized images are captured. The rate at which the images are captured is limited by the frame-rate of the camera sensor – which is in turn limited by the maximum data transfer speed of the connected firewire bus.

In this setup, the laptop 1394 Firewire card is one 800 Mbit/s bus, split into two ports (Hub 0 & Hub 1 of Figure 2.2). Each port has 400 Mbit/s available to it. The image size is 640 $\times$ 480 pixels at 8 bits/pixel. According to the DCAM standard implemented by Point Grey, this equates to a maximum of six cameras being able to run at 15 fps on a single bus [26]. Currently there are five cameras in the system, but this can easily be upgraded to six. Further, because the system is currently only intended for static posture measurement – thus only a single frame is required, the system could capture at a lower frame rate with more cameras per bus. If, however, more cameras are required at a higher frame rate, then more 1394 Firewire adapters must be connected to the computer.

Once the files are captured by the software, they are saved to the computer hard drive and converted from bitmap (*.BMP) format to JPEG (*.JPG) format to reduced file size.

2.5.5 Marker Selection

Currently, markers are selected by hand; the user needs to process each image and select the point as close as possible to the centre of the visible marker (or calibration object landmark). This is a cumbersome and tedious process – especially when there are 41 points to be selected across five images. The improvements to marker selection as well as the reasons for the current method of marker selection are dealt with in §3.3. As was subsequently found, improvements to marker selection play an important role in the accuracy of the reconstruction.

Markers are selected (per image) according to the model definition and saved to a text file in the same order. Where a marker is not visible in the image it is given the value of `'none'`; this is important to keep track of because only markers visible in two or more cameras can be reconstructed to a real world point.

2.5.6 Marker Reconstruction

Triangulation

The marker reconstruction works through the relationship between the image point and the real world point. During calibration, the camera matrix, $\boldsymbol{P}$, is found, which is an homography – mapping of the real world point $\vec{X}$ to the image point (on the image plane) $\vec{x}$, as illustrated in Figure 2.10.

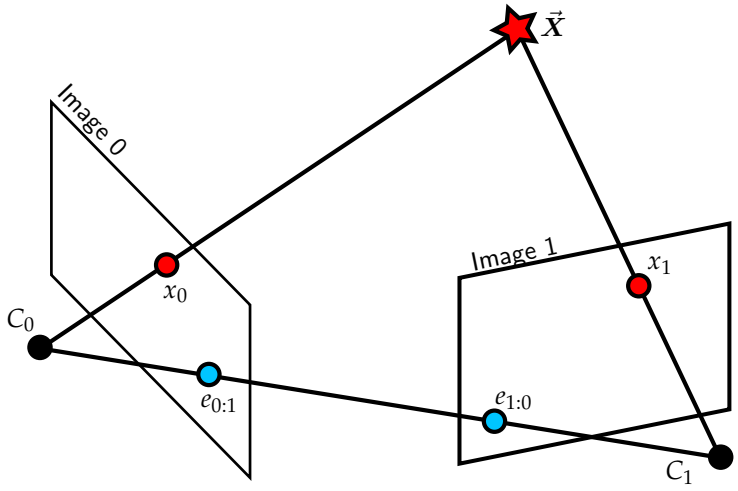


Figure 2.10: The mapping of $\vec{X}$ to two image planes, clearly showing the point correspondence. C_0 and C_1 are the camera centres, $e_{0:1}$ and $e_{1:0}$ are the epipoles (projection of C_1 onto image plane 0), $\vec{x}_0$ and $\vec{x}_1$ are the projected image points, and $\vec{X}$ is the real world point

The point $\vec{X}$ is the unknown real world point. The points $\vec{x}_0$ and $\vec{x}_1$ are corresponding image points, in that they are both the projections of $\vec{X}$ onto different image planes. They relate to the world point via the camera matrix:

$$\vec{x}_0 = \boldsymbol{P}_0 \cdot \vec{X} \tag{2.4}$$

$$\vec{x}_1 = \boldsymbol{P}_1 \cdot \vec{X} \tag{2.5}$$

And, the corresponding points relate to each other through the fundamental matrix, $\boldsymbol{F}$, of the camera pair. This is known as the *epipolar constraint* [3; 2]:

$$\vec{x}_1^{\mathsf{T}} \cdot \boldsymbol{F} \cdot \vec{x}_0 = \emptyset \tag{2.6}$$

Using equations (2.4) and (2.5), as well as the cross-product $\vec{x} \times \boldsymbol{P}\vec{X} = \emptyset$ and the following definitions:

$$\vec{x}_0 = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \in \mathbb{P}^3, \tag{2.7}$$

$$z_0 = 1 = z_1 \quad \text{Projective space constraint,} \tag{2.8}$$

$$\vec{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \in \mathbb{P}^4, \tag{2.9}$$

$$\boldsymbol{P}_0 = \begin{bmatrix} \vec{p}_{0r1} \\ \vec{p}_{0r2} \\ \vec{p}_{0r3} \end{bmatrix}_{3\times 4}, \tag{2.10}$$

a system can be formed and solved, $\boldsymbol{A}\vec{X} = \emptyset$ [27], where:

$$\boldsymbol{A} = \begin{bmatrix} x_0 \cdot \vec{p}_{0r3}^{\mathsf{T}} - z_0 \cdot \vec{p}_{0r1}^{\mathsf{T}} \\ y_0 \cdot \vec{p}_{0r1}^{\mathsf{T}} - z_0 \cdot \vec{p}_{0r2}^{\mathsf{T}} \\ x_1 \cdot \vec{p}_{1r3}^{\mathsf{T}} - z_1 \cdot \vec{p}_{1r1}^{\mathsf{T}} \\ y_1 \cdot \vec{p}_{1r1}^{\mathsf{T}} - z_1 \cdot \vec{p}_{1r2}^{\mathsf{T}} \end{bmatrix} \tag{2.11}$$

The system can be solved using the singular value decomposition (SVD) and direct linear transform (DLT) [27; 2].

Implementation

The marker reconstruction was implemented in the software using the functions in `pyMultiCam` to perform the distortion correction, rectification and DLT triangulation.

The image points and camera matrices per camera are read in from text files. The system must create camera pairs and then check the points in the pairs to ensure that only markers visible in both pairs are triangulated; this is done per

camera pair. Therefore, for five cameras there are ten combinations, thus ten sets of triangulated markers.

The program must then run through the triangulated sets. If a marker appears triangulated in multiple sets, then these duplicate marker values are averaged. For example if marker $\vec{X}_{M3}$ is visible in both camera pairs $C_0 : C_1$ and $C_0 : C_2$, then the final marker value is given by:

$$\vec{X}_{M3} = \frac{\vec{X}_{M3}\Big|_{C_0:C_1} + \vec{X}_{M3}\Big|_{C_0:C_2}}{2} \tag{2.12}$$

After this, if there still are markers remaining that are untriangulated it means that these markers were not visible by at least one camera pair and an error is generated. This process is shown in Figure 2.11.

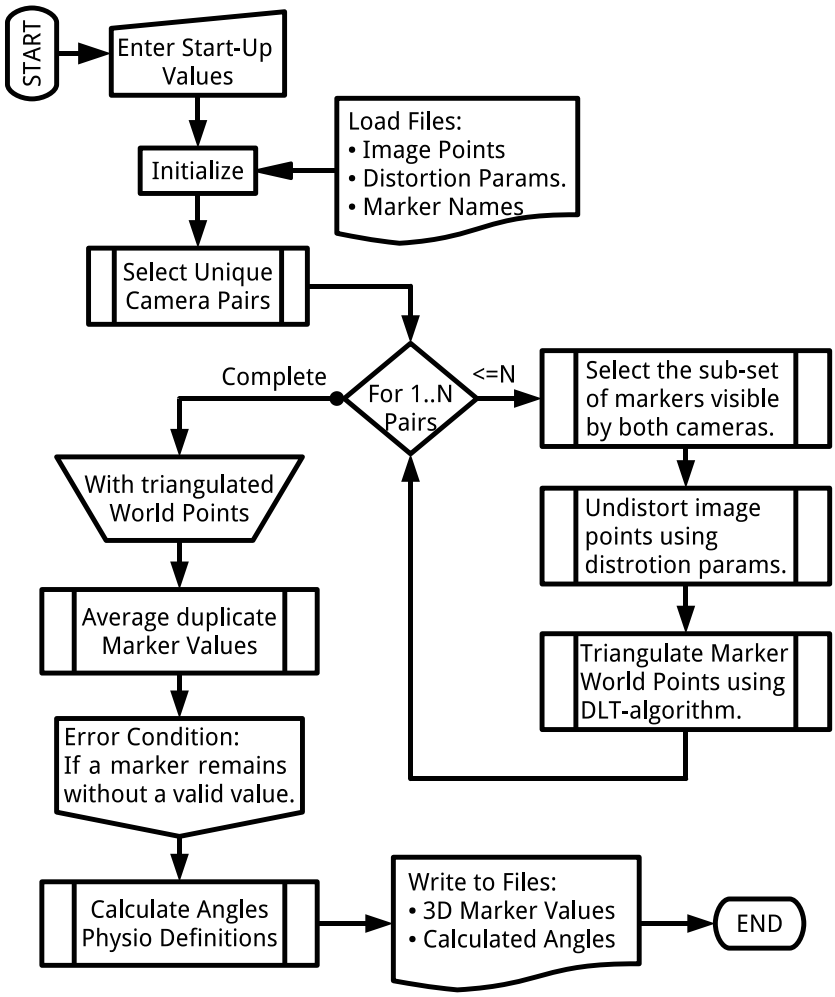


Figure 2.11: Program flow for reconstruction

Once all the markers are available they are written to a text file and the program begins to calculate the required angles.

2.5.7 Angle Calculation

The angles are calculated from the reconstructed world-point values of the markers. Initially, the coordinate system of these markers is the same as that of the calibration object used. During calibration, the object is ideally placed such that the "y-axis" faces in the forward direction of the subject (normal to the *coronal plane*), the "x-axis" faces in the sideways direction from the left to right hip (normal to the *sagittal plane*) and the "z-axis" faces in the upwards direction of gravity (normal to the *transverse plane*); all axes are orthogonal.

To ensure that the markers are indeed body-centred, a series of calculations describe three orthogonal vectors in the calibration reference frame. These describe an axis from the left greater trochanter marker to the right greater trochanter marker as the x-axis. The vertical z-axis is the same in both frames (upwards with gravity) and the forward y-axis is the cross-product of the x and z axes.

Once these three body frame axes are described within the calibration frame, a coordinate frame transformation map is calculated that transforms points in the calibration frame to the body frame.

The reason for this is that if the markers are calculated with an additional machine, the final markers will always be in a coordinate relative to the subject's body. In this way, multiple systems can easily be compared against each other for accuracy.

Once the markers are in the body frame, the angles between specific points are calculated using linear algebra and the dot-product-cosine rule. The mathematics behind angle calculation, and a discussion of error propagation through angle calculations to predict angular accuracy, are provided in §4.1. An example of the definition of one of the calculated angles is shown in Figure 2.12. Further, all angle definitions are found in Appendix §A.

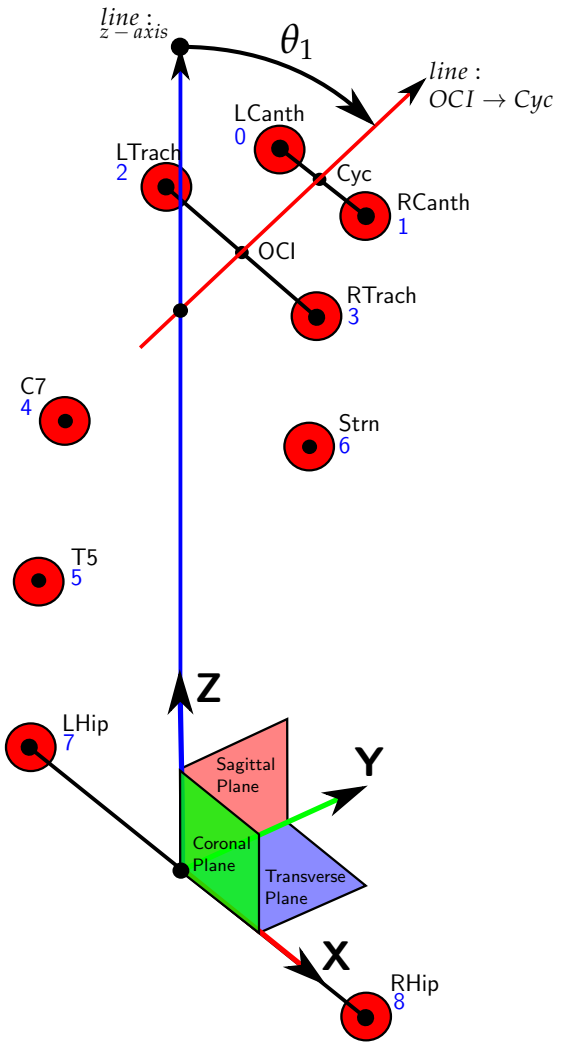


Figure 2.12: Angle definition of head flexion (θ_1) as an example of a typical angle calculated by the system.

Chapter 3

Improvements to the System

3.1 Calibration Ground-Truth Objects

The role of the calibration object is crucial to the working of the system. Even in the case of auto-calibration there needs to be some physical, spatial form from which the metrication parameters can be derived [12]. What follows in §3.1.1 and §3.1.2 is a comparison of the development of a new calibration object to the existing object, to better suit the needs of the system.

3.1.1 Aluminium Corner Object

The initial calibration object was the Aluminium Corner Ground-Truth Calibration Object (Figure 3.1).

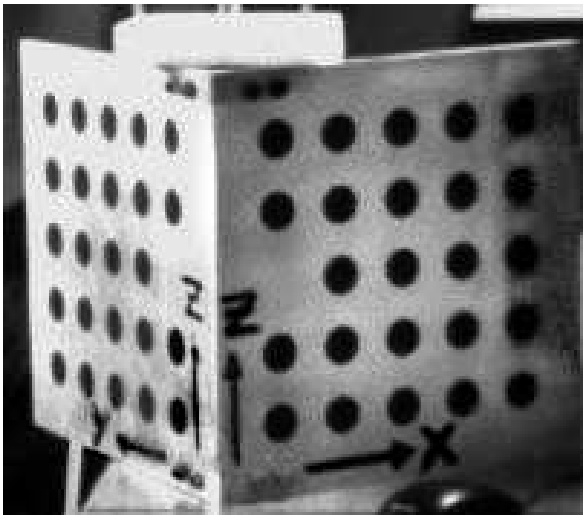


Figure 3.1: Aluminium corner ground-truth calibration object

The object consists of 48 landmarks in two orthogonal planes, XZ and YZ, with the centres of each landmark accurately known. The origin of the object is in the bottommost corner where the two planes meet.

For the application with the PMM, the object was found to have several drawbacks:

- The landmarks are not reflective to infra-red light and as such cannot be used in conjunction with the VICON imaging system for the purposes of verifying the accuracy of the VICON.
- Spatially the object is small, $300 \times 300 \times 300$ mm, severely limiting the capture volume. Ideally, all points to be captured should be inside the calibration area to reduce non-linear errors [2; 28].
- The object is omnidirectional, in that it is best viewed from a head-on direction. In order to capture a subject from a 360° view, the object would need to be rotated so that cameras on the opposite side can see the landmarks. This, however, means that the world origin changes when the object is moved; thus cameras must be paired in sets and some sort of transformation between these two sets must be found.

3.1.2 Pyramid Object

In addressing the problems listed in §3.1.1, a new calibration object had to be designed that would fulfill a role better suited to touring schools and of a size comparable to human subjects from hips to head.

In the end, the *Pyramid* calibration object was created, as seen in Figure 3.2:

Figure 3.2: Pyramid ground-truth calibration object

The object is $700 \times 700 \times 700$ mm, with the reflective spheres ($\varnothing 40$ mm) mounted on wooden dowels; these dowels are securely mounted into a base. The whole object is a wood construction. The markers are arranged on five levels, A to E from lowest to highest, and are named starting from 00, moving counter-clockwise from the $+Y$ axis, as seen in Figure 3.3.

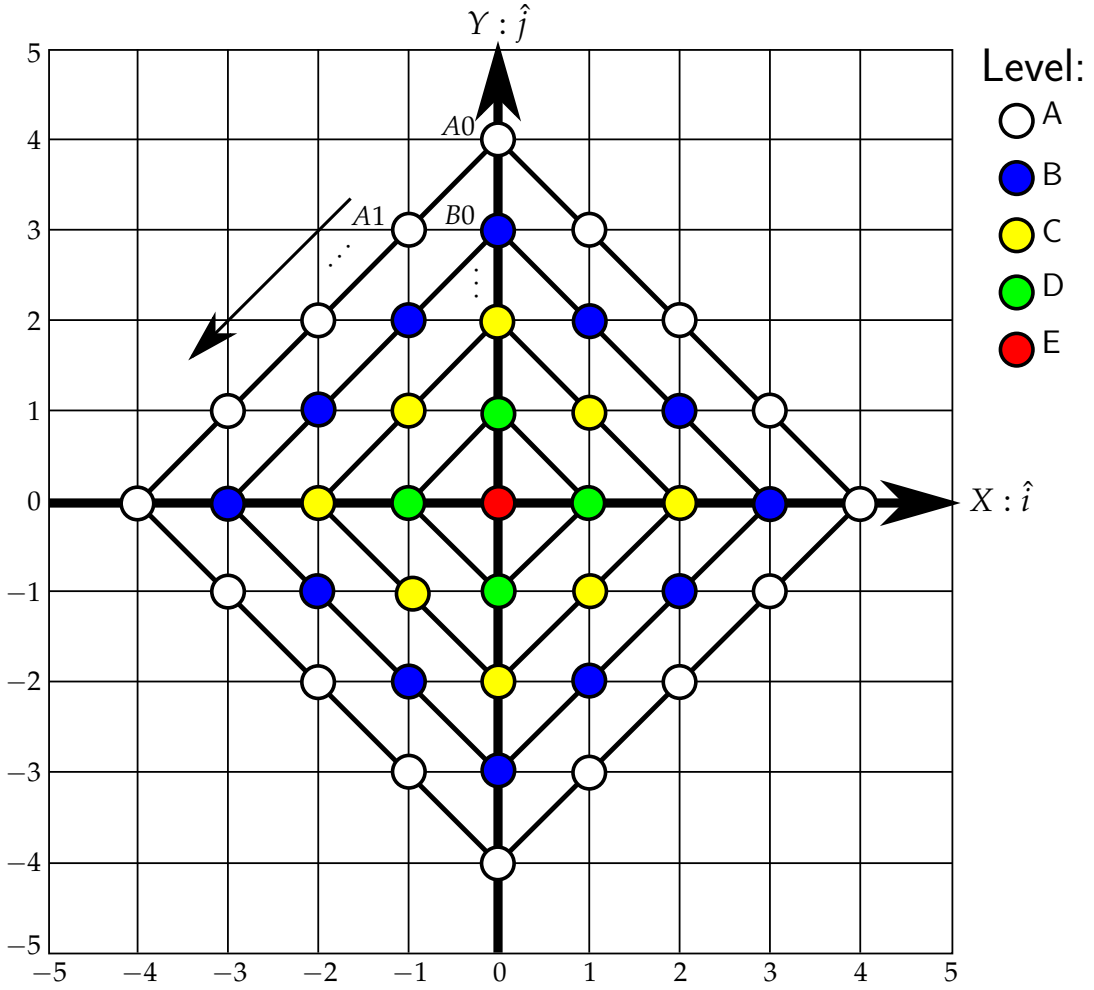


Figure 3.3: Definition and placement of markers for pyramid calibration object

The axes for the system are defined as follows, (right-handed system):

***X*–axis** Line from $A04$ to $A12$.

***Y*–axis** Line perpendicular to X passing through $A00$.

***Z*–axis** Line perpendicular to the plane formed by X and Y, passing through the intersection of X and Y.

The full set of marker coordinates relative to this coordinate reference frame is given in appendix §C.

The main purpose of this object's design (marker configuration) is that cameras positioned all around the object can see enough markers to create an accurate calibration, while not having to move the calibration object for each set of cameras. As a result, all cameras calibrated with this object will have the same defined world coordinate system.

The drawbacks of this configuration are that:

- the object is not as rigid as the aluminium corner object and will break if dropped;
- during CMM measurement, the taller dowels may deflect and invalidate the readings; and
- while the object does fit into a car for transportation, it can be a bit cumbersome to carry around because of its size and fragility. While not as heavy as the corner object, it is still rather leaden.

Despite these negative features of the pyramid object, it was used successfully in the research program, given that:

- it covers the full working volume of a sitting subject;
- the markers are reflective enough to be accurately captured by the VICON system, as discussed in §6.2; and
- all cameras are calibrated to a single world system.

3.2 Lens Distortion Correction

A distortion model is a model that accounts for how the lens of the camera affects images captured by the camera. The linear pinhole camera does not distort the images in any way; light passes through the pinhole onto the image plane in a linear projection. Even when cameras use glass lenses, the assumption of an *ideal lens* can be made, implying that light passing through the lens does so in a way that does not distort the path of the light. However, this assumption is only made to simplify the underlying mathematics of the system – it must be accounted for during implementation.

Once again, a set of assumptions as to the nature of the distortion must be made. There are many models and types of distortion: radial, tangential, decentring and thin-prism being the ones most commonly appearing in the literature. Of these, radial is generally the most important distortion to account for [2]. Accounting for the other models would further improve accuracy in smaller and smaller increments [17].

Further, it must be noted that the signal-to-noise ratio predicts the effectiveness of distortion correction; the larger the signal-to-noise ratio, the more effective the distortion correction algorithms [11]. This obviously means that if there is a lot of noise visible in the images, the cause of the noise must first be removed or corrected for.

Radial distortion can be categorized into three types: barrel, pincushion and complex, usually occurring as either barrel or pincushion, where complex is a combination of the two.

Radial distortion occurs when straight lines in the real world are mapped to curved lines on the image plane. Figure 3.4 shows this effect, where the solid black line is the shape as seen in the image, while the dashed red line is what the shape should look like. The point $\langle x_c, y_c \rangle$ is the distortion centre; all distortion is taken as radial from this point. The red star and blue circle are essentially the same points.

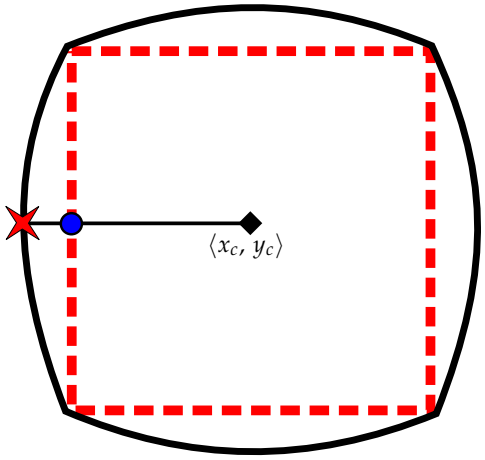


Figure 3.4: Radial distortion

Mathematically, the distorted image point $\langle x, y \rangle$ relates to the true image point (corrected image point) $\langle \hat{x}, \hat{y} \rangle$ through the distance to radial centre r and the distortion factor $L(r)$, as follows [27; 2]:

$$r^2 = (x - x_c)^2 + (y - y_c)^2 \tag{3.1}$$

and

$$L(r) = 1 + \kappa_1 r + \kappa_2 r^2 + \ldots O\left(r^n\right) \tag{3.2}$$

then the corrected points are:

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + L(r) \cdot \begin{bmatrix} x - x_c \\ y - x_c \end{bmatrix} \tag{3.3}$$

Equation (3.2) is typically approximated by an n^{th} order Taylor series expansion [2]. In this case the first two orders are used. This gives rise to the coefficients $\{\kappa_1, \kappa_2\}$. These, together with the distortion centre, form the "distortion parameters" and completely describe the nature of the radial distortion for the given camera at the given calibration – this is important because focal length influences the nature of the distortion and is therefore only applicable to a specific setup.

Because of this, the parameters are solved during image calibration using the calibration object's world and image points to solve the linearizing model, as implemented in pyMultiCam [25]. These parameters are then stored in a text file that accompanies the calibration files. An undistort function can be called up at any time to use the four parameters to undistort a given image point [25].

The results of implementing this improvement are dealt with in §3.4.

3.3 Increasing the Sub-Pixel Accuracy of Marker Selection

3.3.1 Initial Implementation

The most basic method to *digitize* the marker origins, as seen in the images by the cameras, is to visually pick the points by hand. Provided the person performing this is suitably familiar with the process and the model, and remains attentive throughout, spurious marker selection would be minimized.

This method is also the most basic method to implement; simply display the image to the user, loop through the possible markers and allow the user to *pick* the point closest to the origin of the desired marker. In this way the human user would be able to discern the desired marker from the image – as opposed to finding an algorithm that would enable the computer to do this successfully.

However, the process is extremely tedious; during the calibration there are five images with 41 landmarks to be selected per image, giving a total of 205 points that require selection. In the case of the medical study, one data batch required 15 schools of 20 pupils. Calibration for one batch is 3075 points. If there are two follow-up batches, then there are about 9225 points to be selected – purely for calibration and assuming that each click is a perfect selection.

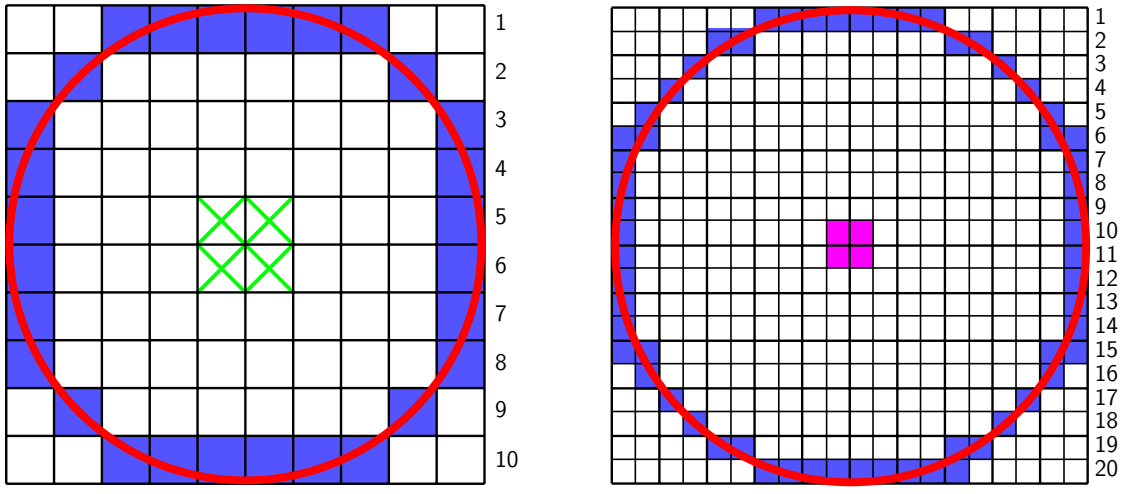
Despite this method being tedious, it was implemented as it forms part of the

minimal machine on which all future improvements can be tested. Furthermore, the input errors are due to human error and rasterization, not due to blob-detection algorithms etc., which fall outside the scope of this design.

3.3.2 Problem with Rasterization

Computers are digital instruments, thus the images are digital representations of what the camera "sees" through the lens. The inside camera sensor (CMOS in this application) is a finite grid of pixel detectors. The incoming light falls upon one of these detectors and triggers a value on the grid. As such the final image is built up of lots of little blocks to represent the light that fell on that detector. If the resolution of the camera is not high enough, there is image degradation in terms of pixellation – where more than one feature in the image is represented by the same pixel.

This type of image is called a bitmap, the grid is a raster and, as is the nature of all things digital, it has quantization effects. Figure 3.5 shows the raster representation of a circle on a 10-pixel grid and a 20-pixel grid; note how increasing the resolution (*pixel depth*) increases the accuracy of the representation of the circle.



(a) Low resolution **(b)** High resolution

Figure 3.5: The rasterization of a circle (red) into a bitmap representation of the circle (blue). The centre of the circle is one of the for possible pixels marked in green.

The biggest problem with rasterization in this application is that the marker centre may not be represented accurately enough. In Figure 3.5, note that the centre

in the low-resolution image is limited to the four pixels with the green crosses, and that they are further from the centre than the four pink pixels in the high-resolution image.

The important, mitigating factor for sub-pixel accuracy is the amount of depth information encoded in the pixel. In Figure 3.6, the real world points project onto the image plane through the camera centre. It can be seen that a change in depth (along the focal axis) only has a small change in the representation on the image plane, whereas a change perpendicular to the focal axis has a more pronounced change on the image plane. Thus the more accurately the change can be represented in the image, the more accurate the reconstruction of real world point from the image data.

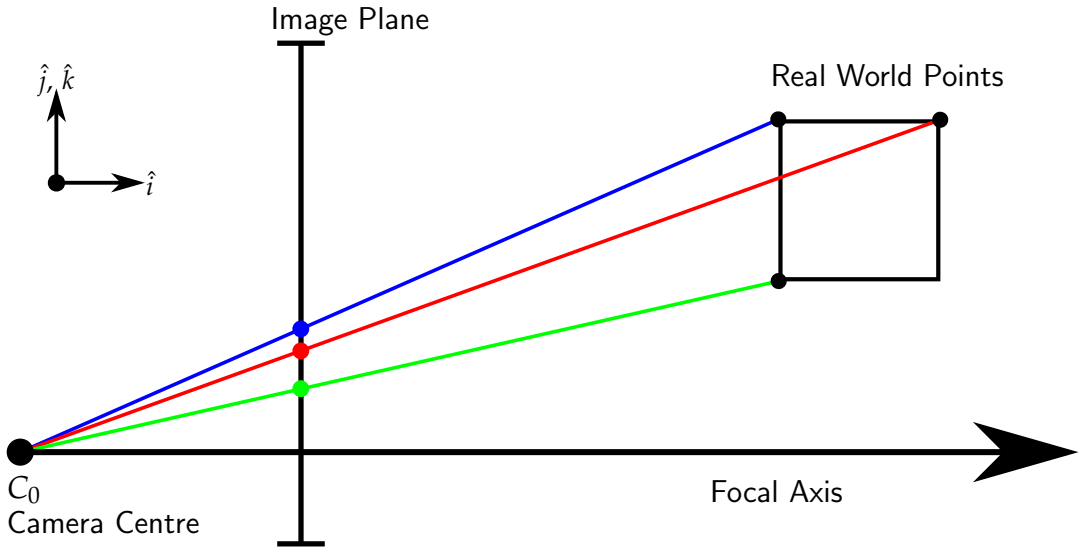


Figure 3.6: 3D points projecting onto image plane, showing how a change in depth along the focal axis has a small change on the image plane. A large amount of depth data is thus encoded in a small amount of image data.

3.3.3 Increasing the Pixel Depth

A higher resolution gives a more accurate digital representation of the image, leading to a more accurate reconstruction. The obvious way to increase the resolution is to use a higher resolution camera/camera setting with tighter focus. However, in this project the image resolution is set at 640×480 px due to camera ability and data bandwidth; with this constraint in place the other way is to artificially increase the

resolution by increasing the pixel depth of the image and then investigate if this increases the accuracy of the system.

Increasing the pixel depth is simply a way of increasing the resolution of the current image, taking a current pixel and dividing it into n smaller pixels of the same colour value as the original pixel. However, this runs the risk of making the image look too jagged. In order to circumvent this the image can be interpolated; while this blurs the image slightly, the objects are more recognizable to the human user than if they are stark blocks [29]. There is an upper limit to how much the pixel depth should be increased by – if increased too far, objects in the image can become indiscernible and therefore useless.

The method to increase pixel depth is similar to zooming an image. The 640×480 px image is displayed in one window, alongside a square 600×600 px zoom window. The user selects a square potion (size s_{width}) of the image around the marker they would like to select. This portion is then scaled to fit into the zoomed window – effectively increasing the region to a depth of 600×600 px; the zoomed window is interpolated using bi-cubic interpolation.

In the zoomed window, the user uses a yellow circle with cross hairs to best fit the circle around the marker in such a way that the centre of the cross-hair in the yellow circle matches the centre of the marker – this centre is the centre of the marker in the zoomed window and must first be scaled back correctly to the original image window's pixel scale.

Figure 3.7 shows the process of selecting the zoom area (green square in A), then fitting the yellow circle in B. Finally, the pink circle in C shows how the yellow circle fits over the original image.

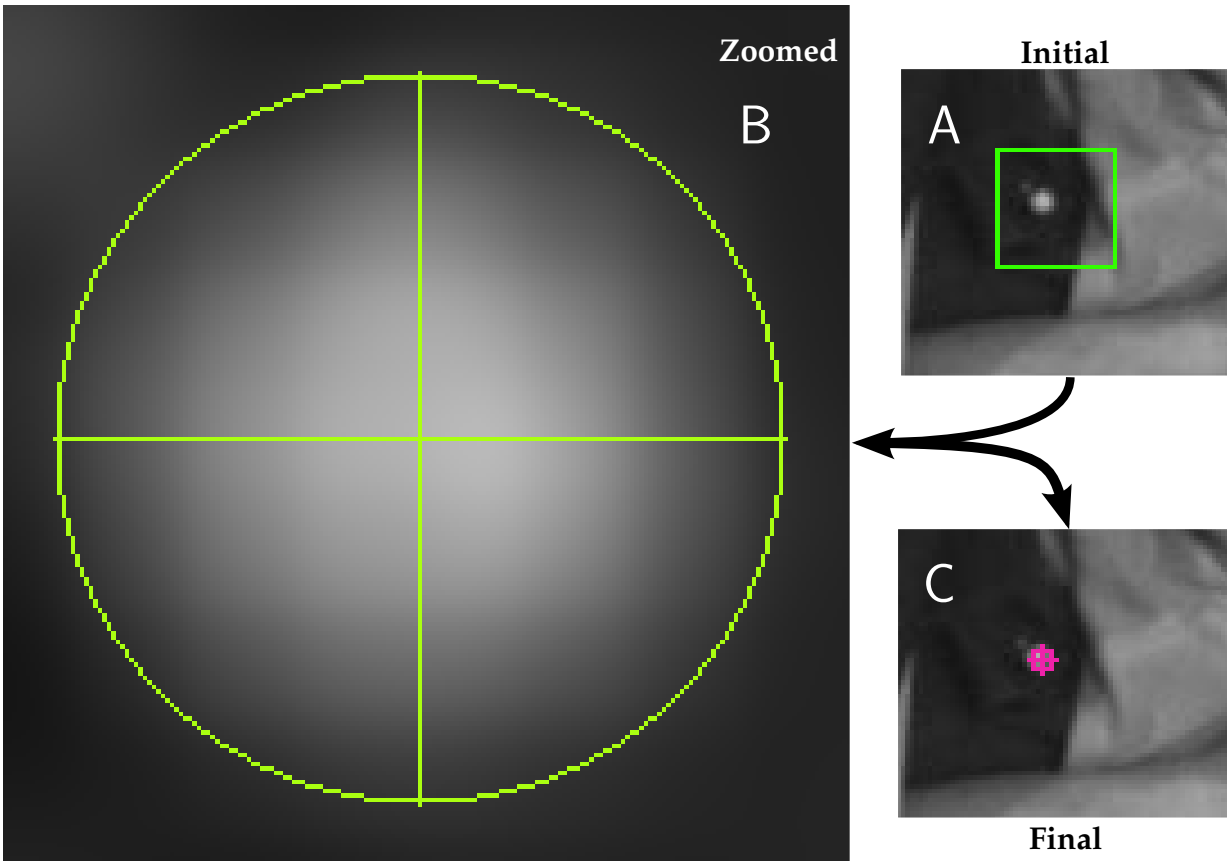


Figure 3.7: Increasing pixel depth and marker selection

It is important to clarify that, because there is a change in scale of the "zoom" region, it is not possible to simply take the pixel value of the marker centre in the zoom window as the pixel value of the marker in the image. Figure 3.8 illustrates how the zoom window and zoom region relate to the image window. The desired image point is the point $\langle x_c, y_c \rangle$ in the image window, and the point selected is $\langle x_{zc}, y_{zc} \rangle$ in the zoom window.

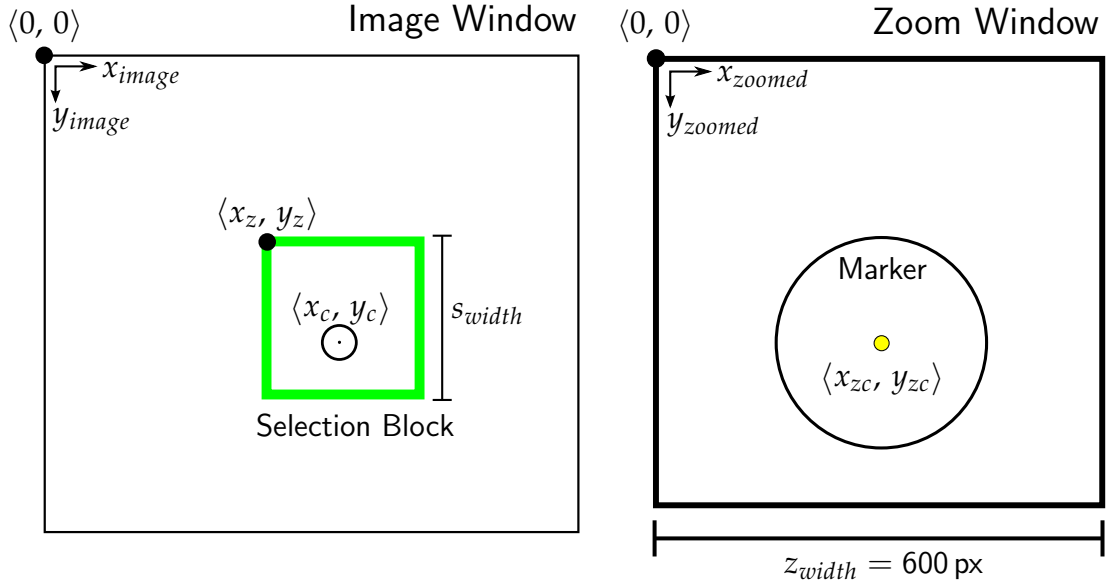


Figure 3.8: Zooming principle for marker selection

Given the starting coordinates of the zoom region in the image window as $\langle x_z, y_z \rangle$, with a region width of s_{width} and a zoom window width of z_{width}, a scaling factor f_{sz} can be defined that scales between the image and zoom windows. Note that selection region and zoom window in this application specifically are square so that there is only one scale factor and not a scale factor for each direction.

The desired marker centre in the image window can now be described as:

$$\begin{bmatrix} x_c \\ y_c \end{bmatrix} = \begin{bmatrix} x_z \\ y_z \end{bmatrix} + f_{sz} \cdot \begin{bmatrix} x_{zc} \\ y_{zc} \end{bmatrix} \tag{3.4}$$

where:

$$f_{sz} = \frac{s_{width}}{z_{width}} \tag{3.5}$$

The results of these improvements are discussed in §3.4.

There is a problem that relates to pixel accuracy that must be noted: The PMM captures its images in bitmap format and then converts them to JPEG format – a compression format that greatly reduces the file size (on average from 900 kB to 40 kB in this application).

The problem is that the JPEG compression algorithm is a lossy algorithm – there is data degradation in the image. Unfortunately this was only realized at the time of writing up this thesis and, as a result, the effects of the JPEG compression on the accuracy have not been quantified.

It is theorized that the effects should be small, as there is no decrease in pixel depth, and that the total degradation should not exceed 8% of the total image quality as specified as an input to the compression algorithm.

Upon realizing this problem, the best suggestion would be to quantify the compression effects on the accuracy and, in future, convert all BMP bitmap images to PNG format (a lossless compression image format).

It also means that all improvement accuracies shown in §3.4 may be improved further by moving away from the JPEG format.

3.4 End Results of the Improvements

A method had to be defined to verify that these suggested improvements were actual improvements, and to see to what extent these improvements improve the system.

Ideally, the reconstructed data would be compared to real world known data and, based on that comparison, the improvements would be quantified. However, at the time of this part of the investigation the new pyramid calibration object was only being constructed and the only available data was what had been captured for the physiotherapists using the aluminium corner object, inside which there is no standard metric beyond the calibration object.

Fortunately the markers used are all of a standard size, either ∅9.5 mm, ∅10 mm, ∅12.5 mm or ∅15 mm, depending on what is available to the physiotherapists. It must be noted that all markers used on a single subject for a capture must all be of the same size; different sized markers should not be used concurrently. The standard size for the markers meant that, if a small cluster of points was selected inside a single marker in all available image views and then reconstructed, the distances between all the points in the reconstructed cluster should be smaller or equal to the diameter of the marker used – as all the points are selected from inside the marker.

Therefore, given a set of points $\{\vec{x}_n\} \mapsto \{\vec{X}_n\}$, selected specifically from inside the visible marker, the resulting 3D reconduction points must meet the following criteria:

$$\|\vec{X}_i - \vec{X}_j\| \leq \varnothing D_{marker}, \forall\, i \neq j \in n \tag{3.6}$$

To achieve this with the PMM, five images were selected and in those images a marker was chosen. Inside that specific marker twelve points were selected in the

image, according to the pattern shown in figure 3.9:

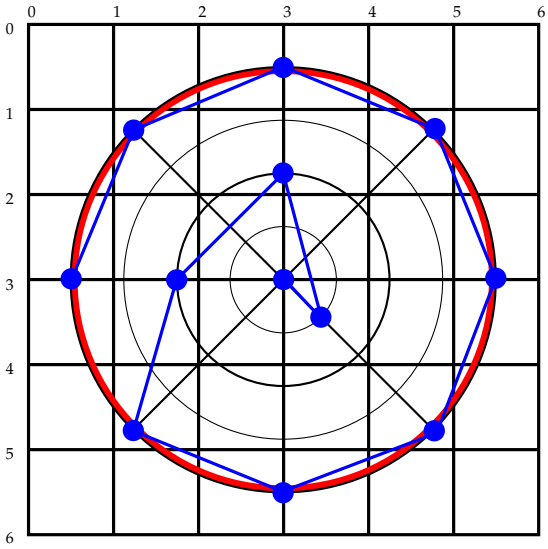
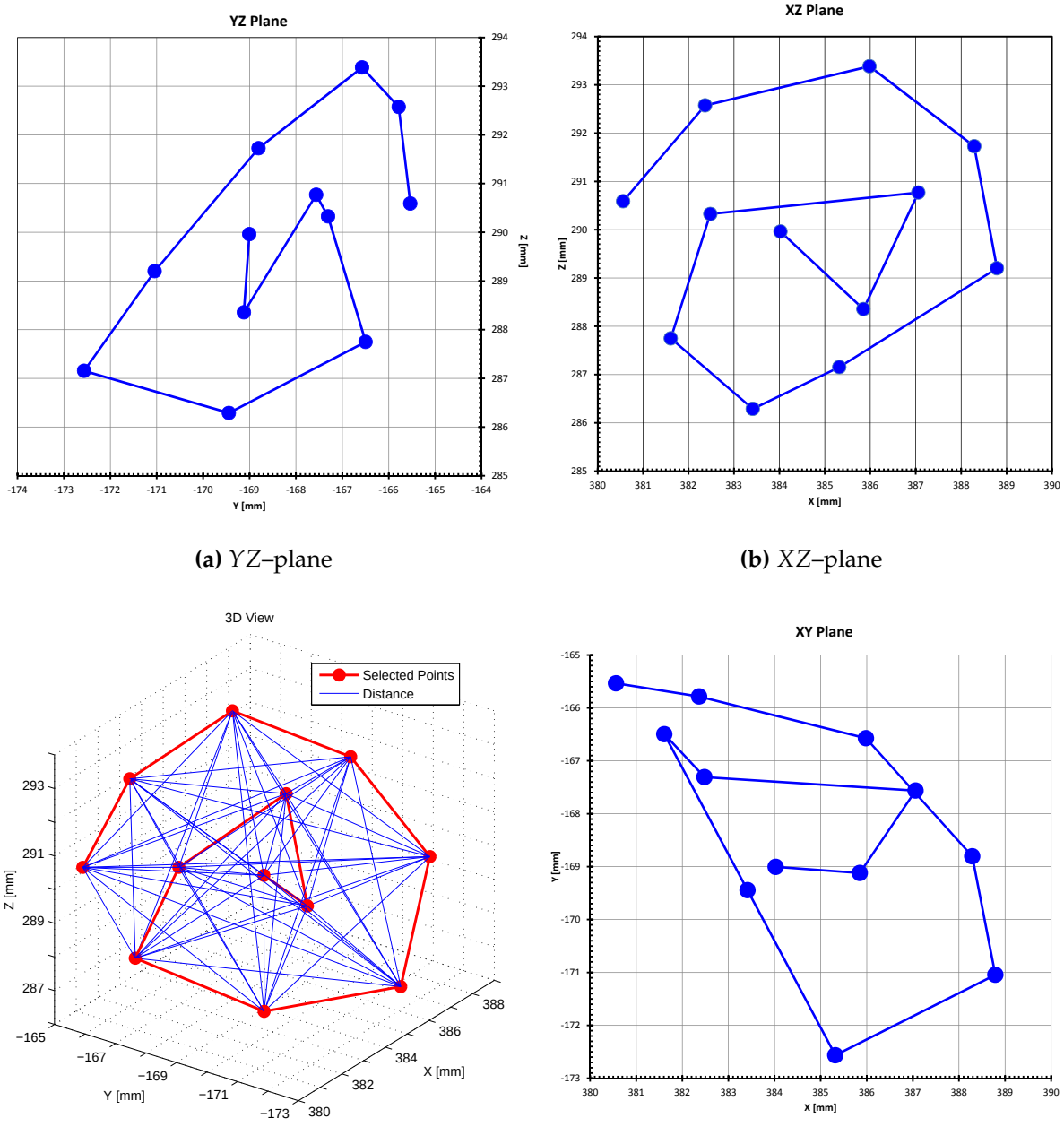


Figure 3.9: Pattern for selecting points within a marker; the red circle represents the marker boundary and the blue points are the selected points within the marker. The distances between these points is used to determine accuracy improvements. The grid illustrates a 1×1 px raster.

Figure 3.10 shows the 3D reconstructed points selected on a ∅10 mm marker, with full distortion correction and sub-pixel accuracy implemented. Figure 3.10c is the 3D view and shows the distances between the selected points.



(a) YZ–plane

(b) XZ–plane

(c) $3D$–view

(d) XY–plane

Figure 3.10: Reconstruction from points selected within a ∅10 mm marker; grid spacing represents 1 mm intervals.

In carrying out this investigation, the points in the image were selected to correspond as close as possible and to be as consistent as possible if re-selection was required. The images were processed under four different conditions:

Nothing — Neither distortion correction nor sub-pixel accuracy enhancement implemented; absolute base level.

Distortion — Only distortion correction implemented.

Sub-Pixel — Only sub-pixel accuracy implemented (required re-selection).

Final — Both distortion correction and sub-pixel accuracy enhancement implemented.

For each condition the raw data is reconstructed to form a data set of 3D reconstructed points. The relative distance was calculated between each point and every other point in the data set, and because these points were all selected within a ∅10 mm marker the distances calculated from these points should be lesser than or equal to the diameter of the marker, equation (3.6).

Figure 3.11 illustrates the data as generated by applying equation (3.6) to the reconstructed data sets for each condition. The plot is semi-log to illustrate how vastly each set differs.

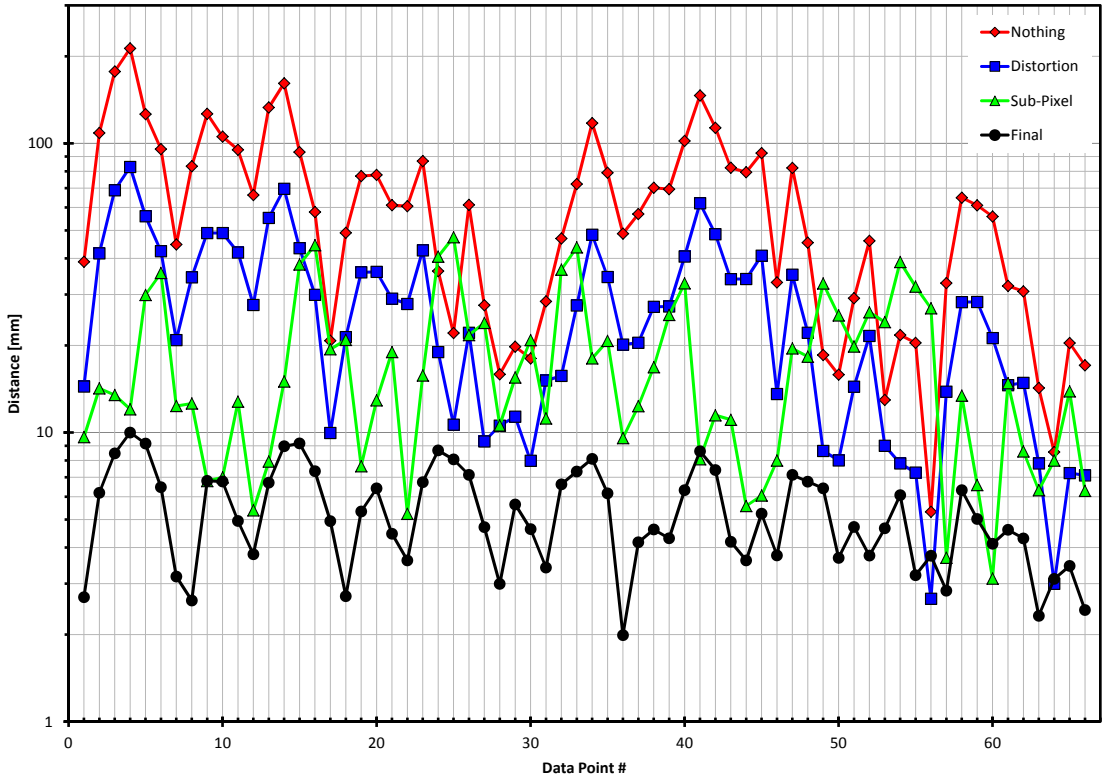


Figure 3.11: Distance between points selected within a marker under different conditions; each data set represent is independent and illustrates overall system performance for a given condition.

Each data set presented should be considered independent and points in one set should not be directly compared to points in another set. This is due to the need to *re-select* the points within images for the purposes of implementing the sub-pixel accuracy improvement. As a result the points may not directly correspond between sets.

Histograms of the data seem to indicate that the data is not Gaussian distributed; however, the nature of the distribution of such data is outside the scope of this thesis and needs to be investigated further. As a result, the average values and standard deviations shown in Figure 3.12 serve as a metric to gauge the improvement of each initial suggestion, as well as the overall combined result.

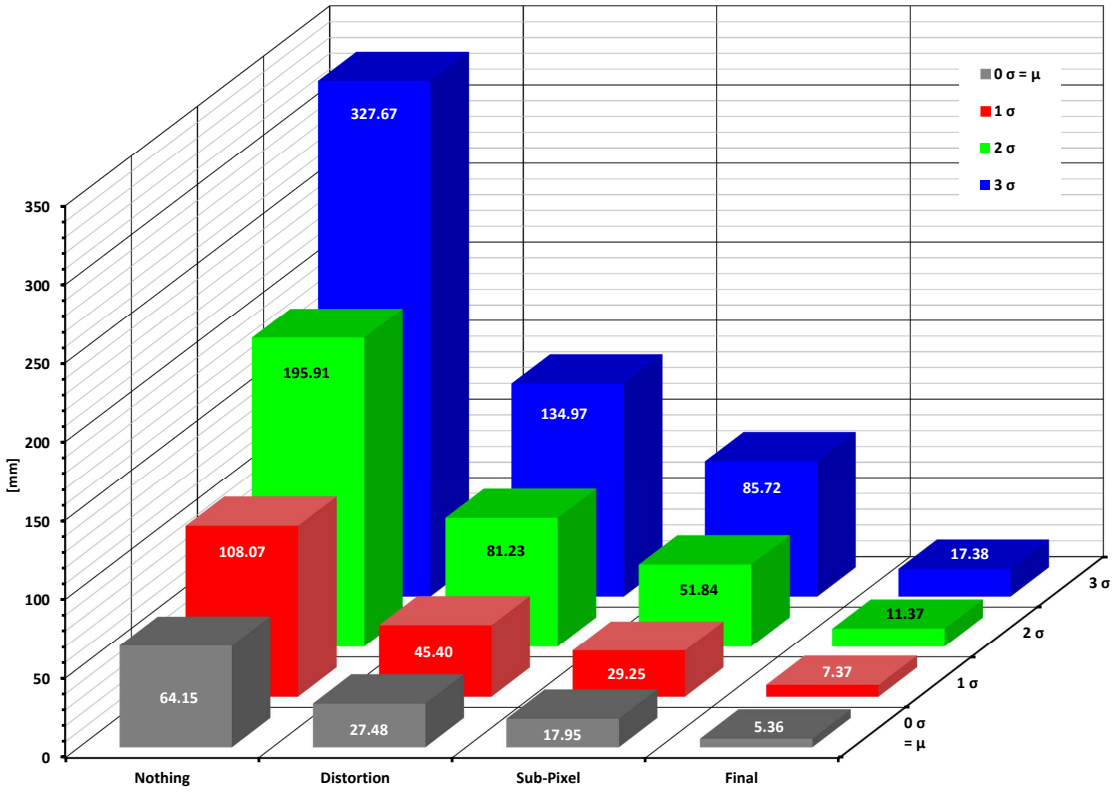


Figure 3.12: Visual representation of statistical data for each case

Each initial suggestion, distortion correction and sub-pixel accuracy enhancement offers substantial improvement, individually, from the initial, *base* values. However, it is when the two are implemented in conjunction that the full value of the improvements is harnessed. Table 3.1 shows the average value for a set and the percentage improvement of the average from the base set, *Nothing*.

Table 3.1: Statistical data for each condition to test improvements

	Nothing	**Distortion**	**Sub-Pixel**	**Final**
Average	64.15 mm	27.48 mm	17.95 mm	5.36 mm
Improvement		57.16%	72.01%	91.64%

This improvement directly reflects the true meaning of the data that was presented in figure 3.11. For each presented condition the markers should ideally be bounded by 10 mm which is clearly not the case in until both improvements are fully implemented.

Chapter 4

Error Propagation from Points to Angles

Once points in an image have been reconstructed to points in 3D space they can be used in further applications. As noted in the application of the physiotherapy study, the markers are used to determine posture angles.

As with any real world system there is inherent uncertainty attached to all aspects that compose the overall system. It is important that users and designers understand the effects that these uncertainties have upon the workings of the system, in order to better account for them, mitigate their effects and, where possible, minimize these effects.

The purpose of this chapter is to investigate the effects of the marker position uncertainty in 3D space, propagating through the mathematics, to affect the final angle value calculated for a set of markers. It is not to determine or derive the positional uncertainty of the reconstructed marker positions in $\mathbb{R}^3$ as a result of the underlying uncertainties in the triangulation, pixel selection and camera model.

4.1 Background Mathematics

Given three points in $\mathbb{R}^3$:

$$\begin{aligned} \vec{A} &= \langle a_x, a_y, a_z \rangle \\ \vec{B} &= \langle b_x, b_y, b_z \rangle \\ \vec{C} &= \langle c_x, c_y, c_z \rangle \end{aligned}$$

The desired angle is between $A\widehat{C}B$, and this can be described by two vectors, one from $\vec{C}$ to $\vec{A}$ and one from $\vec{C}$ to $\vec{B}$. Then the angle θ_{ACB} can be found using the dot-product.

$$\vec{U} = \vec{A} - \vec{C} \tag{4.1}$$

$$\vec{V} = \vec{B} - \vec{C} \tag{4.2}$$

then:

$$\theta_{ACB} = \arccos\left(\frac{\vec{U}\cdot\vec{V}}{\left\|\vec{U}\right\|\cdot\left\|\vec{V}\right\|}\right) \tag{4.3}$$

As can been seen, the angle is a function of the two direction vectors; the two direction vectors, in turn, are functions of the three spatial points, and these are ultimately described by three variables in x, y, z. This means that the angle θ_{ACB} is ultimately a function of nine variables – an important consideration to bear in mind when calculating the effects of the uncertainty of each of those variables upon the final angle.

4.2 The Effects of the Uncertainties of x, y, z upon θ

Given the nature of the stereographic reconstruction, uncertainties can be the result of many underlying factors. A significant cause can be that the non-linear nature of the camera exceeds the linear model of the camera and that correction algorithms are not 100% effective.

Another important factor is that three-dimensional data has been encoded into two-dimensional data by way of the camera. The result of this is that the uncertainties are more significant along one axis (the camera's principle axis). Large amounts of spatial data have been encoded into a small amount of camera data. The effect is that small errors in selecting the correct point in the camera data will result in a large error in the resulting spatial reconstruction. Spatial changes in the horizontal and vertical planes are easier to record on the camera plane than changes in depth. This can be mitigated by camera placement and number of cameras in the system, but the important aspect that it highlights is that uncertainties in x, y, z for a given point will not necessarily be the same, i.e. in general, $\sigma_x \neq \sigma_y \neq \sigma_z$. This means that, while the angle is a function of nine variables, each variable has its own uncertainty acting upon it.

So, here is a function or system that takes the nine input variables and performs operations on them to produce an output $Y = F(X)$. It has been found that if X has a covariance matrix Λ_X, then the covariance of the result can be given by *the error propagation law* [30]:

$$\boldsymbol{\Lambda_Y} = \boldsymbol{J_F}(\bar{\boldsymbol{X}}) \cdot \boldsymbol{\Lambda_X} \cdot \boldsymbol{J_F}(\bar{\boldsymbol{X}})^{\mathsf{T}}, \tag{4.4}$$

where $\boldsymbol{J_F}$ is the Jacobian Matrix (Jacobian) of the function F, and $\bar{X}$ is the mean value as an estimation of the true X within a sample group.

By having an analytical solution of the Jacobian Matrix for the angle function given by equation (4.3), the effects of the uncertainties on point positions can be used to described the resulting angle uncertainties.

4.3 Deriving the Jacobian Matrix for an Angle Between Two Vectors in $\mathbb{R}^3$

The Jacobian is defined as the matrix of first-order partial-derivatives of a vector or scalar function $F(X)$ with respect to another vector. This means that the partial derivative for θ_{ACB} must be found for $\langle A, B, C \rangle$.

In deriving the derivatives, the function can be going to be broken into smaller sub-functions so that differentiation rules like the chain rule, quotient rule, etc. may be applied and clarity in the logic may be achieved. The argument vector X consists of the three elements of each of the point vectors, A, B, C, and can be written as:

$$\begin{aligned} \vec{X} &= \langle a_x, a_y, a_z, b_x, b_y, b_z, c_x, c_y, c_z \rangle \\ &= \left\langle \vec{A}, \vec{B}, \vec{C} \right\rangle \end{aligned} \tag{4.5}$$

Taking

$$\theta(\vec{X}) = \arccos\left(\frac{\vec{U} \cdot \vec{V}}{\left\|\vec{U}\right\| \cdot \left\|\vec{V}\right\|} \right) \tag{4.6}$$

let

$$\begin{aligned} F(\vec{X}) &= \vec{U} \cdot \vec{V} \\ &= u_x \cdot v_x + u_y \cdot v_y + u_z \cdot v_z \end{aligned} \tag{4.7}$$

$$\begin{aligned} G(\vec{X}) &= \left\| \vec{U} \right\| \\ &= \sqrt{\left(u_x^2 + u_y^2 + u_z^2 \right)} \end{aligned} \tag{4.8}$$

$$\begin{aligned} H(\vec{X}) &= \left\| \vec{V} \right\| \\ &= \sqrt{\left(v_x^2 + v_y^2 + v_z^2 \right)} \end{aligned} \tag{4.9}$$

$$\begin{aligned} \therefore \theta(\vec{X}) &= \arccos\left(\frac{F(\vec{X})}{G(\vec{X}) \cdot H(\vec{X})} \right) \\ E(\vec{X}) &= \frac{F(\vec{X})}{G(\vec{X}) \cdot H(\vec{X})} \\ \Rightarrow \theta\left(\vec{X}\right) &= \arccos\left(E(\vec{X}) \right) \end{aligned} \tag{4.10}$$

Now the partial derivative with respect to some variable $x \in \vec{X}$ can be written as:

$$\begin{aligned} \frac{\partial}{\partial x} \theta(\vec{X}) &= \frac{\partial \arccos\left(E(X)\right)}{\partial E(\vec{X})} \cdot \frac{\partial E(\vec{\vec{X}})}{\partial x} \\ &= \left(\frac{-1}{\sqrt{(1 - E(X)^2)}} \right) \cdot \frac{\partial}{\partial x} E(\vec{X}) \end{aligned} \tag{4.11}$$

Using the quotient and product rules:

$$\frac{\partial}{\partial x} E(\vec{X}) = \tag{4.12}$$

$$\frac{\left(\frac{\partial}{\partial x} F(\vec{X}) \right) G(\vec{X}) H(\vec{X}) - F(\vec{X}) \cdot \left(\frac{\partial}{\partial x} G(\vec{X}) \right) H(\vec{X}) - F(\vec{X}) G(\vec{X}) \cdot \left(\frac{\partial}{\partial x} H(\vec{X}) \right)}{G(\vec{X})^2 \cdot H(\vec{X})^2}$$

A deeper look at $F(\vec{X})$ shows:

$$\begin{aligned} F(\vec{X}) &= \vec{U} \cdot \vec{V} \\ &= \left(\vec{A} - \vec{C} \right) \cdot \left(\vec{B} - \vec{C} \right) \\ &= \left(a_x - c_x \right) \left(b_x - c_x \right) + \left(a_y - c_y \right) \left(b_y - c_y \right) + \left(a_z - c_z \right) \left(b_z - c_z \right) \\ &= \left(a_x b_x + a_y b_y + a_z b_z \right) - \left(a_x c_x + a_y c_y + a_z c_z \right) \ldots \\ &\quad - \left(b_x c_x + b_y c_y + b_z c_z \right) + \left(c_x^2 + c_y^2 + c_z^2 \right) \end{aligned} \tag{4.13}$$

Table 4.1 shows the partial derivatives of $F(\vec{X})$ per element and how they can bet written in a condensed vector form:

Table 4.1: Partial derivatives of $F(\vec{X})$, equation (4.13)

$\frac{\partial}{\partial a_x} F(\vec{X}) = b_x - c_x$	$\frac{\partial}{\partial b_x} F(\vec{X}) = a_x - c_x$	$\frac{\partial}{\partial c_x} F(\vec{X}) = -a_x - b_x + 2c_x$
$\frac{\partial}{\partial a_y} F(\vec{X}) = b_y - c_y$	$\frac{\partial}{\partial b_y} F(\vec{X}) = a_y - c_y$	$\frac{\partial}{\partial c_y} F(\vec{X}) = -a_y - b_y + 2c_y$
$\frac{\partial}{\partial a_z} F(\vec{X}) = b_z - c_z$	$\frac{\partial}{\partial b_z} F(\vec{X}) = a_z - c_z$	$\frac{\partial}{\partial c_z} F(\vec{X}) = -a_z - b_z + 2c_z$
$\Rightarrow \frac{\partial}{\partial \vec{A}} F(\vec{X}) = \vec{V}$	$\frac{\partial}{\partial \vec{B}} F(\vec{X}) = \vec{U}$	$\frac{\partial}{\partial \vec{C}} F(\vec{X}) = -\left(\vec{U} + \vec{V}\right)$

Now the derivative of $F(\vec{X})$ may be written in vector form:

$$\begin{aligned} \frac{\partial}{\partial \vec{X}} F(\vec{X}) &= \left\langle \frac{\partial}{\partial \vec{A}} F(\vec{X}), \frac{\partial}{\partial \vec{X}} F(\vec{X}), \frac{\partial}{\partial \vec{X}} F(\vec{X}) \right\rangle \\ &= \left\langle \vec{V}, \vec{U}, -\left(\vec{U} + \vec{V}\right) \right\rangle \end{aligned} \tag{4.14}$$

With $G(\vec{X})$ it follows that

$$\begin{aligned} G(\vec{X}) &= \sqrt{\left(u_x^2 + u_y^2 + u_z^2\right)} \\ &= \sqrt{\left((a_x - c_x)^2 + (a_y - c_y)^2 + (a_z - c_z)^2\right)} \\ &= \sqrt{\left(a_x^2 + a_y^2 + a_z^2 - 2a_x c_x - 2a_y c_y - 2a_z c_z + c_x^2 + c_y^2 + c_z^2\right)} \end{aligned} \tag{4.15}$$

Using

$$\frac{\mathrm{d}}{\mathrm{d}\,x} \sqrt{S(x)} = \frac{1}{2\sqrt{S(x)}} \cdot \left(\frac{\mathrm{d}}{\mathrm{d}\,x} S(x)\right)$$

Here, Table 4.2 shows the partial derivatives of $G(\vec{X})$ per element and then written in a condensed vector form:

Table 4.2: Partial derivatives of $G(\vec{X})$, equation (4.15)

$\frac{\partial}{\partial a_x} G(\vec{X}) =$	$\frac{a_x - c_x}{G(\vec{X})}$	$\frac{\partial}{\partial b_x} G(\vec{X}) =$	0	$\frac{\partial}{\partial c_x} G(\vec{X}) =$	$\frac{c_x - a_x}{G(\vec{X})}$
$\frac{\partial}{\partial a_y} G(\vec{X}) =$	$\frac{a_y - c_y}{G(\vec{X})}$	$\frac{\partial}{\partial b_y} G(\vec{X}) =$	0	$\frac{\partial}{\partial c_y} G(\vec{X}) =$	$\frac{c_y - a_y}{G(\vec{X})}$
$\frac{\partial}{\partial a_z} G(\vec{X}) =$	$\frac{a_z - c_z}{G(\vec{X})}$	$\frac{\partial}{\partial b_z} G(\vec{X}) =$	0	$\frac{\partial}{\partial c_z} G(\vec{X}) =$	$\frac{c_z - a_z}{G(\vec{X})}$
$\Rightarrow \frac{\partial}{\partial \vec{A}} G(\vec{X}) =$	$\frac{\vec{U}}{\|U\|}$	$\frac{\partial}{\partial \vec{B}} G(\vec{X}) =$	$\emptyset$	$\frac{\partial}{\partial \vec{C}} G(\vec{X}) =$	$-\frac{\vec{U}}{\|U\|}$

The derivative of $G(\vec{X})$ in vector form becomes

$$\frac{\partial}{\partial \vec{X}} G(\vec{X}) = \left\langle \frac{\vec{U}}{\|U\|}, \emptyset, -\frac{\vec{U}}{\|U\|} \right\rangle \tag{4.16}$$

Similarly, with $H(\vec{X})$:

$$\begin{aligned} H(\vec{X}) &= \sqrt{\left(v_x^2 + v_y^2 + v_z^2\right)} \\ &= \sqrt{\left((b_x - c_x)^2 + (b_y - c_y)^2 + (b_z - c_z)^2\right)} \\ &= \sqrt{\left(b_x^2 + b_y^2 + b_z^2 - 2b_x c_x - 2b_y c_y - 2b_z c_z + c_x^2 + c_y^2 + c_z^2\right)} \end{aligned} \tag{4.17}$$

Table 4.3 shows the partial derivatives of $H(\vec{X})$ per element and the resulting vector form:

Table 4.3: Partial derivatives of $H(\vec{X})$, equation (4.17).

$\frac{\partial}{\partial a_x} H(\vec{X}) =$	0	$\frac{\partial}{\partial b_x} H(\vec{X}) =$	$\frac{b_x - c_x}{H(\vec{X})}$	$\frac{\partial}{\partial c_x} H(\vec{X}) =$	$\frac{c_x - b_x}{H(\vec{X})}$
$\frac{\partial}{\partial a_y} H(\vec{X}) =$	0	$\frac{\partial}{\partial b_y} H(\vec{X}) =$	$\frac{b_y - c_y}{H(\vec{X})}$	$\frac{\partial}{\partial c_y} H(\vec{X}) =$	$\frac{c_y - b_y}{H(\vec{X})}$
$\frac{\partial}{\partial a_z} H(\vec{X}) =$	0	$\frac{\partial}{\partial b_z} H(\vec{X}) =$	$\frac{b_z - c_z}{H(\vec{X})}$	$\frac{\partial}{\partial c_z} H(\vec{X}) =$	$\frac{c_z - b_z}{H(\vec{X})}$
$\Rightarrow \frac{\partial}{\partial \vec{A}} H(\vec{X}) =$	$\emptyset$	$\frac{\partial}{\partial \vec{B}} H(\vec{X}) =$	$\frac{\vec{V}}{\|\vec{V}\|}$	$\frac{\partial}{\partial \vec{C}} H(\vec{X}) =$	$-\frac{\vec{V}}{\|\vec{V}\|}$

With the derivative of $H(\vec{X})$ in vector form, as

$$\frac{\partial}{\partial \vec{X}} H(\vec{X}) = \left\langle \emptyset, \frac{\vec{V}}{\left\|\vec{V}\right\|}, -\frac{\vec{V}}{\left\|\vec{V}\right\|} \right\rangle \tag{4.18}$$

The partial derivative of $E(\vec{X})$ can now be written as

$$\frac{\partial}{\partial \vec{X}} E(\vec{X}) \;=\; \dots$$

$$\frac{\begin{bmatrix} \vec{V} \\ \vec{U} \\ -\left(\vec{U}+\vec{V}\right) \end{bmatrix} \left\|\vec{U}\right\| \cdot \left\|\vec{V}\right\| - \left(\vec{U}\cdot\vec{V}\right) \begin{bmatrix} \frac{\vec{U}}{\|\vec{U}\|} \\ \emptyset \\ -\frac{\vec{U}}{\|\vec{U}\|} \end{bmatrix} \left\|\vec{V}\right\| - \left(\vec{U}\cdot\vec{V}\right) \begin{bmatrix} \emptyset \\ \frac{\vec{V}}{\|\vec{V}\|} \\ -\frac{\vec{V}}{\|\vec{V}\|} \end{bmatrix} \left\|\vec{V}\right\|}{\left\|\vec{U}\right\|^2 \cdot \left\|\vec{V}\right\|^2}$$

$$= \begin{bmatrix} \frac{1}{\|\vec{U}\|\cdot\|\vec{V}\|}\cdot\vec{V} - \frac{(\vec{U}\cdot\vec{V})}{\|\vec{U}\|^3\cdot\|\vec{V}\|}\cdot\vec{U} \\ \frac{1}{\|\vec{U}\|\cdot\|\vec{V}\|}\cdot\vec{U} - \frac{(\vec{U}\cdot\vec{V})}{\|\vec{U}\|\cdot\|\vec{V}\|^3}\cdot\vec{V} \\ \left(\frac{(\vec{U}\cdot\vec{V})}{\|\vec{U}\|^3\cdot\|\vec{V}\|} - \frac{1}{\|\vec{U}\|\cdot\|\vec{V}\|}\right)\cdot\vec{U} + \left(\frac{(\vec{U}\cdot\vec{V})}{\|\vec{U}\|\cdot\|\vec{V}\|^3} - \frac{1}{\|\vec{U}\|\cdot\|\vec{V}\|}\right)\cdot\vec{V} \end{bmatrix} \tag{4.19}$$

For interest's sake, it can be noted that $\frac{\partial}{\partial \vec{C}} E(\vec{X}) = -\frac{\partial}{\partial \vec{A}} E(\vec{X}) - \frac{\partial}{\partial \vec{B}} E(\vec{X})$.

Finally, the Jacobian of $\theta_{ABC}(\vec{X})$ will have the form,

$$\boldsymbol{J}_{\boldsymbol{\theta}_{\boldsymbol{ACB}}}(\vec{X}) = \left\langle \; \frac{\partial\,\theta_{ACB}(\vec{X})}{\partial\, a_x}, \;\; \dots, \;\; \frac{\partial\,\theta_{ACB}(\vec{X})}{\partial\, c_z} \; \right\rangle, \tag{4.20}$$

and can be written as

$$\vec{\boldsymbol{J}}_{\boldsymbol{\theta}_{\boldsymbol{ABC}}}(\vec{X}) = \dots \tag{4.21}$$

$$\left(\frac{-1}{\sqrt{1-\left(\frac{\vec{U}\cdot\vec{V}}{\|\vec{U}\|\cdot\|\vec{V}\|}\right)^2}}\right) \cdot \begin{bmatrix} \frac{1}{\|\vec{U}\|\cdot\|\vec{V}\|}\cdot\vec{V} - \frac{(\vec{U}\cdot\vec{V})}{\|\vec{U}\|^3\cdot\|\vec{V}\|}\cdot\vec{U} \\ \frac{1}{\|\vec{U}\|\cdot\|\vec{V}\|}\cdot\vec{U} - \frac{(\vec{U}\cdot\vec{V})}{\|\vec{U}\|\cdot\|\vec{V}\|^3}\cdot\vec{V} \\ \left(\frac{(\vec{U}\cdot\vec{V})}{\|\vec{U}\|^3\cdot\|\vec{V}\|} - \frac{1}{\|\vec{U}\|\cdot\|\vec{V}\|}\right)\cdot\vec{U} + \left(\frac{(\vec{U}\cdot\vec{V})}{\|\vec{U}\|\cdot\|\vec{V}\|^3} - \frac{1}{\|\vec{U}\|\cdot\|\vec{V}\|}\right)\cdot\vec{V} \end{bmatrix}$$

Further, substituting in the unit-vectors, $\hat{\boldsymbol{u}} = \frac{\vec{U}}{\|\vec{U}\|}$ and $\hat{\boldsymbol{v}} = \frac{\vec{V}}{\|\vec{V}\|}$ into equation (4.3) and re-arranging, the Jacobian thus becomes:

$$\vec{\boldsymbol{J}}_{\boldsymbol{\theta}_{\boldsymbol{ABC}}}(\vec{X}) = \dots \tag{4.22}$$

$$\left(\frac{-1}{\sqrt{1-(\hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{v}})^2}}\right) \cdot \begin{bmatrix} \frac{1}{\|\vec{U}\|}\left(\hat{\boldsymbol{v}} - (\hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{v}})\cdot\hat{\boldsymbol{u}}\right) \\ \frac{1}{\|\vec{V}\|}\left(\hat{\boldsymbol{u}} - (\hat{\boldsymbol{v}}\cdot\hat{\boldsymbol{u}})\cdot\hat{\boldsymbol{v}}\right) \\ \frac{-1}{\|\vec{U}\|}\left(\hat{\boldsymbol{v}} - (\hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{v}})\cdot\hat{\boldsymbol{u}}\right) + \frac{-1}{\|\vec{V}\|}\left(\hat{\boldsymbol{u}} - (\hat{\boldsymbol{v}}\cdot\hat{\boldsymbol{u}})\cdot\hat{\boldsymbol{v}}\right) \end{bmatrix}$$

4.4 Finding the Input Covariance

Recall the general system, $Y = F(X)$; with the law of error propagation it was said that the covariance of Y, (Λ_Y), relates to the covariance of X, (Λ_X), through the Jacobian of F. Mathematically the law is written as $\boldsymbol{\Lambda_Y} = \boldsymbol{J_F}(\bar{\boldsymbol{X}}) \cdot \boldsymbol{\Lambda_X} \cdot \boldsymbol{J_F}(\bar{\boldsymbol{X}})^{\mathsf{T}}$, from equation (4.4).

However, the Jacobian of the function is only one of two components of the system. The other piece is equally important and is the actual covariance matrix of X, which describes how accurate the marker points are in 3D.

It is here that the calibration data becomes important. During calibration, the image points of the calibration object were picked out of the images and used with their corresponding ground truth world points to calculate the camera matrix.

These image points can now be used with the camera matrices to determine the real world points of the calibration object. The calculated world points can then be compared to the ground truth world points (measured on the computer measuring machine (CMM)). From this data the standard deviation of the point errors can be calculated – these would then be the values that form the covariance matrix.

The covariance matrix changes per calibration of the machine, as it is dependent on the camera matrices. How this is implemented in the PMM, as well as how to find the covariance matrix from an example data set, is discussed in §6.4.2.

Because there are three spatial dimensions, the covariance matrix will have the form:

$$\boldsymbol{\Lambda_{world}} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy}^2 & \sigma_{xz}^2 \\ \sigma_{xy}^2 & \sigma_y^2 & \sigma_{yz}^2 \\ \sigma_{xz}^2 & \sigma_{yz}^2 & \sigma_z^2 \end{bmatrix} \tag{4.23}$$

However, because $\boldsymbol{J_F}(\bar{\boldsymbol{X}})$ has nine variables, $x - y - z$ for each of the three points in the world, the covariance matrix will have the form:

$$\boldsymbol{\Lambda_{\vec{X}}} = \begin{bmatrix} \boldsymbol{\Lambda_{world}} & \emptyset & \emptyset \\ \emptyset & \boldsymbol{\Lambda_{world}} & \emptyset \\ \emptyset & \emptyset & \boldsymbol{\Lambda_{world}} \end{bmatrix}_{9\times 9} \tag{4.24}$$

4.5 Verifying the Error Model

There is no point in having a model that is not actually correct; to this end the model must be verified. Considering that the model predicts the angular error for

an angle calculated between three points in $\mathbb{R}^3$, the best way to test it would be to create three points, create an error distribution around those points of known deviation, and then calculate the angles for all deviated points to get an angular data set. At the same time, the original true points and standard deviations with the error model must be used to predict the angular deviation. This is then compared to the statistical deviation of the data set, and these two values should be the same. Figure 4.1 defines three points in space; the angle θ is calculated between two vectors, $\vec{U}$ and $\vec{V}$, as described previously in §4.1. The point A' is a disturbed point and illustrates how this would affect the desired angle. Where the angle to be measured is between $A'CB$, the angle would be $\theta_{ACB} + \delta\theta$.

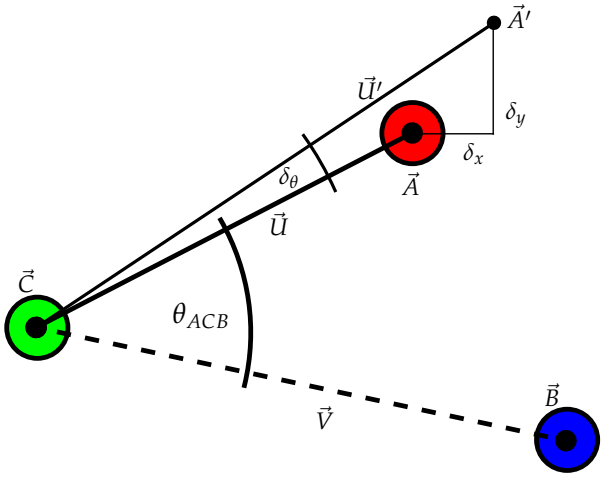


Figure 4.1: Definition of angle θ_{ACB} between three points ABC showing offset errors on point A resulting in an error offset angle

Thus, to test the accuracy, a specific set of points of a known distribution had to be created. Figure 4.2 illustrates this, whereby these points have a zero mean and a 5 mm standard deviation in all axes. The value of 5 mm was chose because it is half the diameter of the markers currently being used, thus 10 mm, the marker diameter has 95% confidence interval (95% of all reconstructions of the marker will have an error of less than 10 mm).

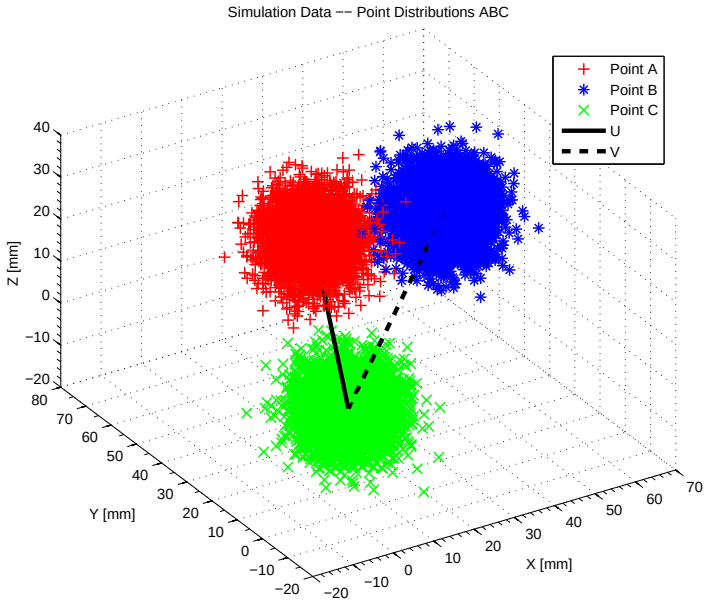


Figure 4.2: Generated simulation data gaussian estimations of points ABC: $\sigma_x = \sigma_y = \sigma_z = 5\,\text{mm}$

However, the deviations in the axes are not all the same [15; 10]. Using the PMM Mannequin Test data from Table 6.12 in §6.4, Figure 4.3 illustrates the effects:

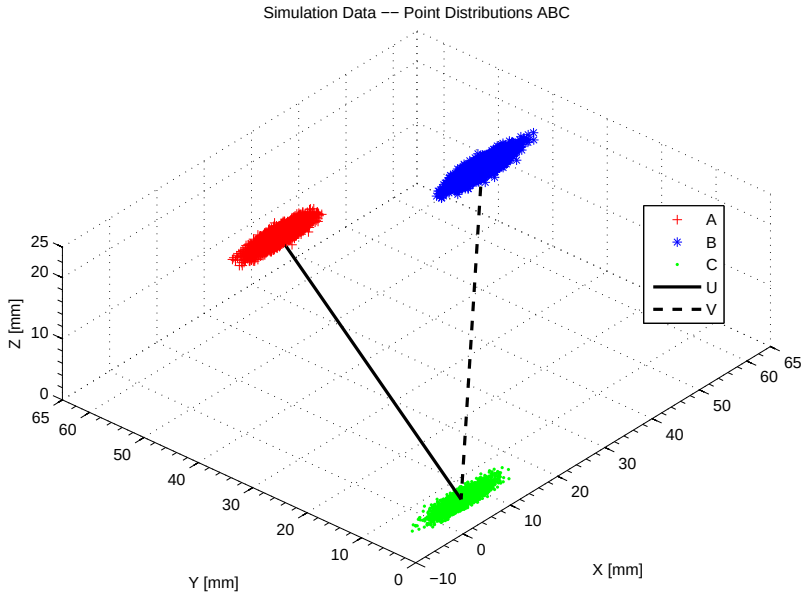


Figure 4.3: Generated simulation data gaussian estimations of points ABC based on PMM data:
$\sigma_x = 2.65\,\text{mm}$, $\sigma_y = 0.55\,\text{mm}$, $\sigma_z = 0.56\,\text{mm}$

A point was taken from each point cloud. Using these three points, the angle was calculated and stored until all points in the cloud had been selected. From the set of angles the standard deviation was calculated, σ_{data}. At the same time, the standard deviations used to generate the point clouds were used with equation 4.3 to calculate the predicted angular deviation, σ_{model}; the difference between σ_{data} and σ_{model} would show how accurate the model is. Figure 4.4 illustrates the angle calculation for the data set of the PMM. In the figure the angle θ_{ACB} is plotted for all the points in ABC. The average angle, $\mu = 27.8°$, is drawn in with both the 1σ interval for σ_{data} and σ_{model}. The error between the model estimate and the data value of the standard deviation was $\epsilon_{prediction} = 0.02°$.

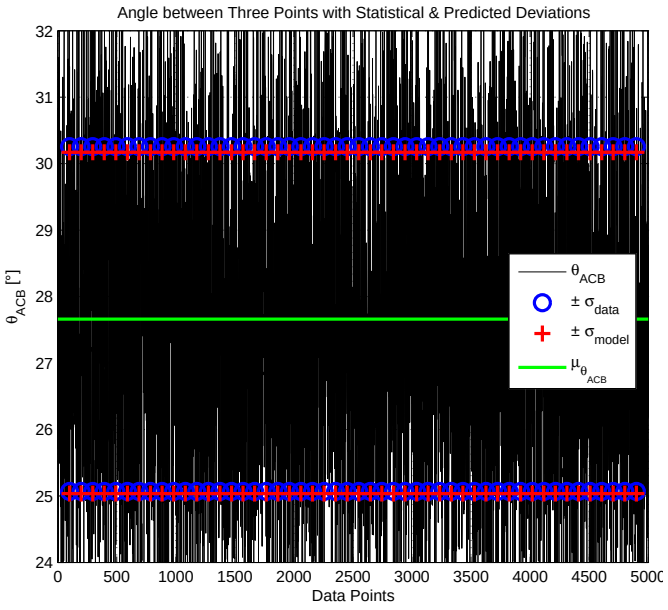


Figure 4.4: Calculated θ_{ABC} from data set showing statistical mean & standard deviation and model predicted standard deviation.

From the angles calculated, a histogram was created, which can be seen in Figure 4.5. This shows that angles calculated from points that are Gaussian distributed are themselves Gaussian distributions.

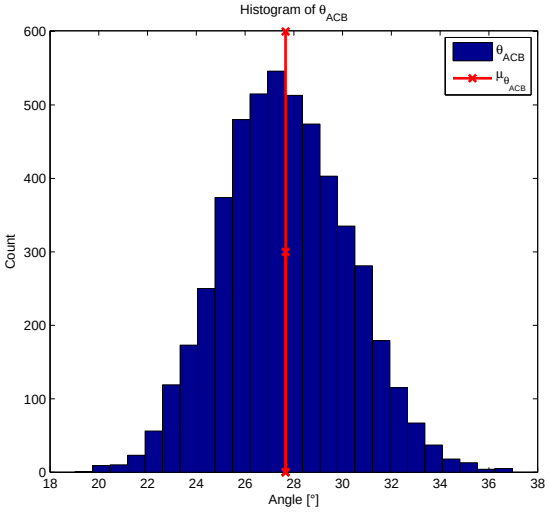


Figure 4.5: Histogram of the angle between three points ABC

Further testing showed that, for the same data, *Poisson Distributed*, the resulting angular error seems to remain Gaussian; this was not verified, however, and requires further investigation. The prediction error increases to around $\epsilon_{prediction} =$

0.2° for Poisson-distributed marker positions. Beyond one clinical set of data, the points A, B, and C were randomized in space. They were then disturbed by a randomized deviation. These randomized deviations were further used to predict the angular error. The randomized distributions were used to calculate the angular deviation. The whole process essentially randomizes the data seen in Figure 4.3. This was then repeated multiple times, and the prediction error between model and data for all these cases was calculated. The statistical results for the prediction error comparing the error model to actual data is shown in Table 4.4. This randomization meant that short and long vectors would result, as well as wide and narrow vectors in varying orientations – basically covering as much of feasible $\mathbb{R}^3$ as computationally possible.

Table 4.4: Statistical data for angular error model prediction error

	$\epsilon_{prediction}$
μ_{Total}	$-0.0004°$
σ_{Total}	$0.0390°$
Points per Cloud	1000
Predictions	1000

As can be seen, the model is an accurate predictor of the expected angular error for given deviations in the marker positions.

4.6 Error Model Implications

The model is mathematically correct and appropriately accurate, for the machine user this implies: If the distances between markers are known, and the system's uncertainty is known, then the expected uncertainty can be calculated. Thus, if an experiment is to be designed using the PMM with distances between markers in the range of 50 mm to 500 mm and uncertainties of $\sigma_x = 2.65\,\text{mm}$, $\sigma_y = 0.55\,\text{mm}$, $\sigma_z = 0.56\,\text{mm}$, then the expected angular uncertainty, $\sigma_{\theta_{ACB}}$, can be calculated across this range, as in Figure 4.6:

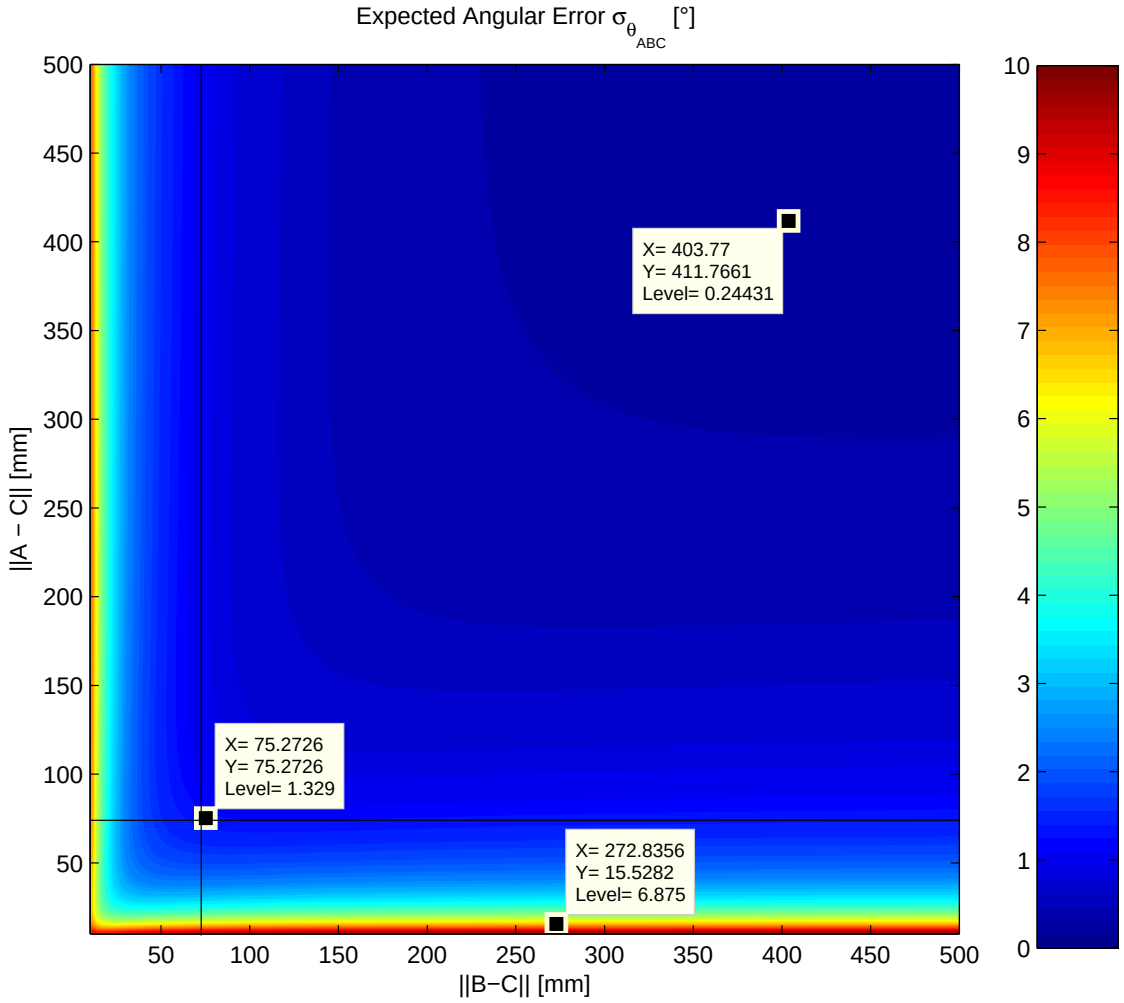


Figure 4.6: Predicted angular deviation as a function of distance between points

Note how quickly the error falls off as $||U||$ and $||V||$ increase; to this end, Figure 4.7 is the same data plotted on logarithmic axes. Here it can be seen clearly that the plot is not symmetrical; this is because the deviations along the different axes are not the same. Due to the fact that the points A and B may be extending more in one axis than another, they will be experiencing more of that axis's deviation.

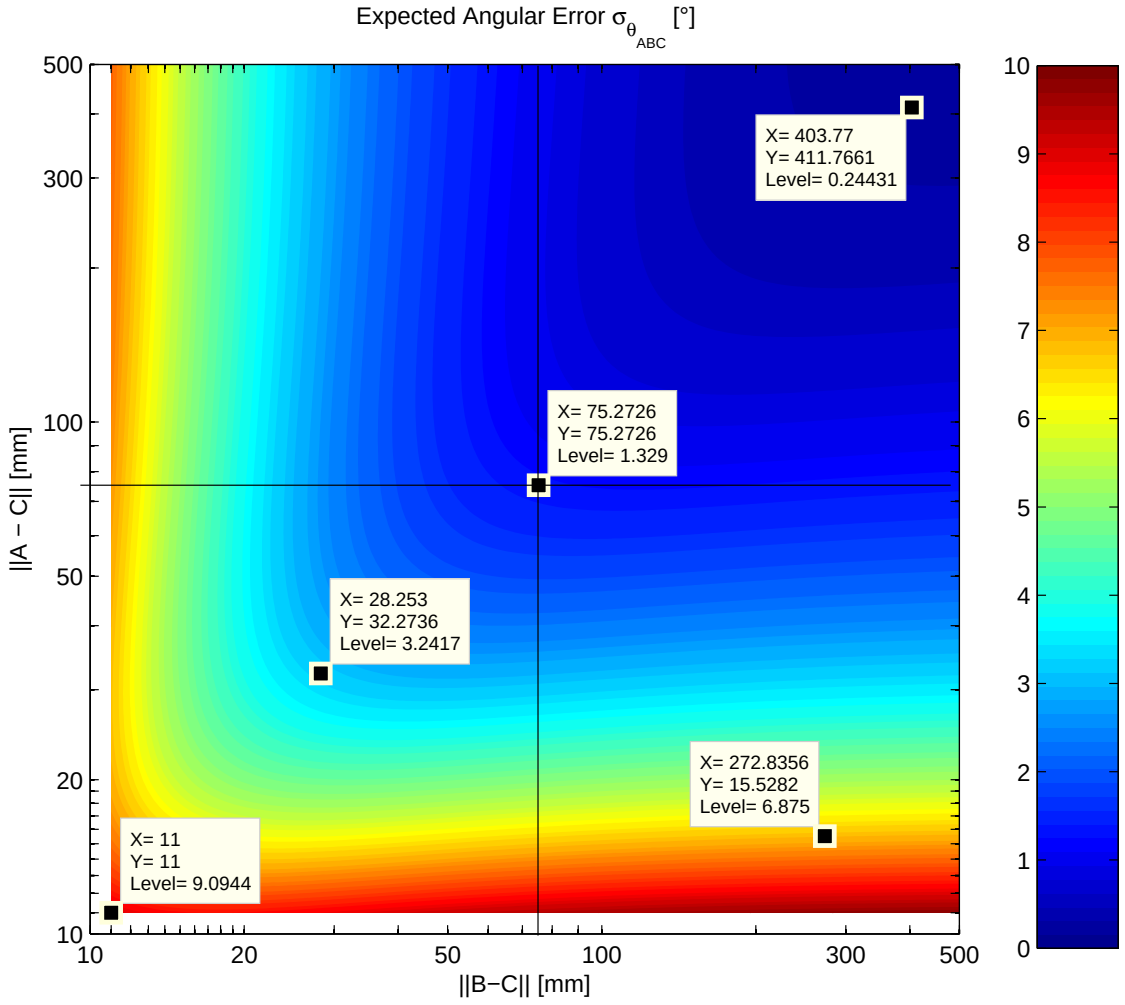


Figure 4.7: Predicted angular deviation as a function of distance between points (logarithmically scaled)

At the 75 mm point (per example, distance between Author's ear and eye – shortest distance between markers in the Physiotherapy study), the predicted angular error is $\sigma_\theta = 1.329°$, which is below the acceptable 5° value and well below the absolute worst value of 10°.

To illustrate how different input deviations for the markers affect the angular deviation, a manifold plot is drawn in Figure 4.8. Each surface corresponds to an angular deviation error σ_θ for the three inputs: $||U||$, $||V||$ and σ_{xyz}. Note that for this simulation the vectors extend in the $\langle 1, \pm 1, 1 \rangle$ directions, so that effects of deviations are not biased by direction. The deviations for each axis are equal, $\sigma_x = \sigma_y = \sigma_z = \sigma_{xyz}$ in millimetres. The manifold surfaces correspond to all points in $||U||$, $||V||$ and σ_{xyz} that give rise to a specific value of $\sigma\theta_{ACB}$.

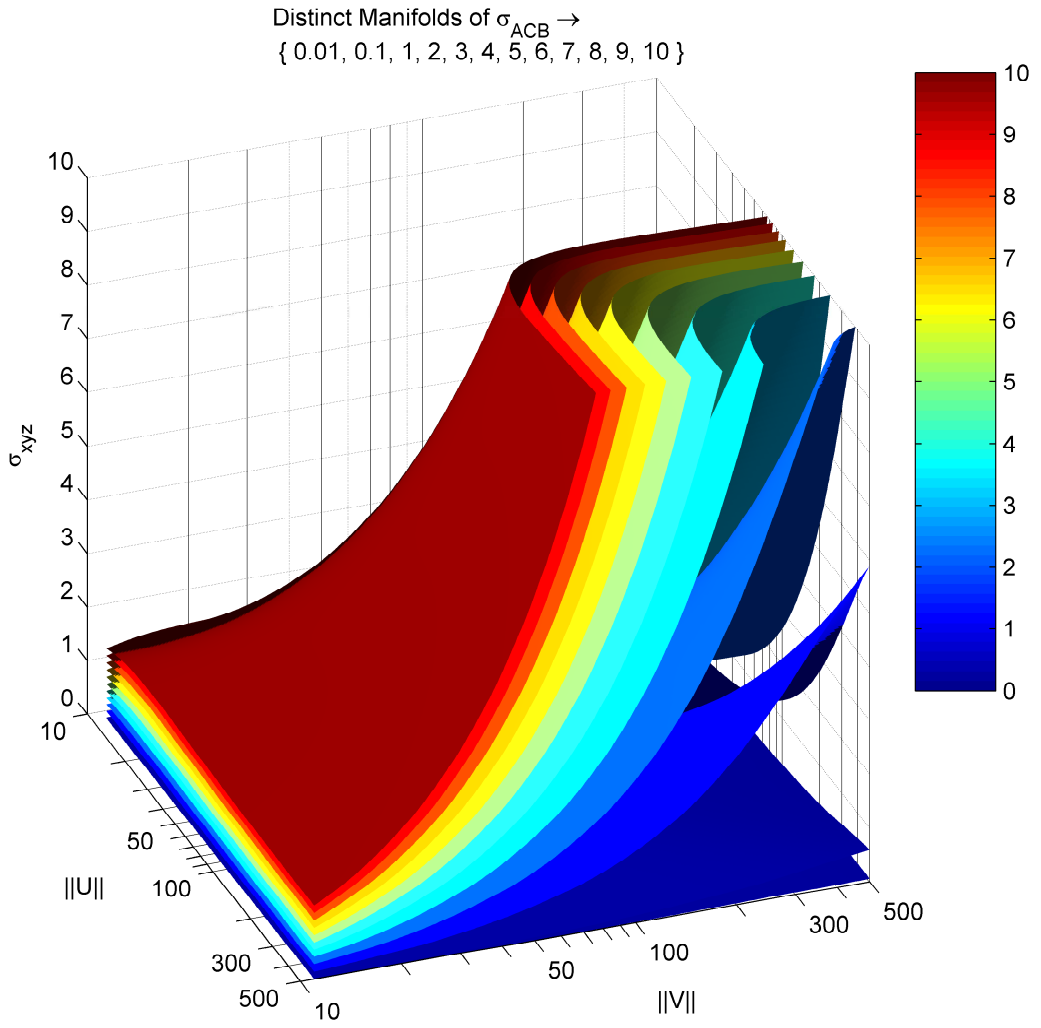


Figure 4.8: Distinct manifolds of $\sigma_{\theta_{ACB}} = \{0.01, 0.1, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}[^\circ]$

To illustrate the effects of the spatial input deviations, σ_x, σ_y, σ_z, on σ_θ, Figure 4.9 was generated for the points $||U|| = ||V|| = 100\,\text{mm}$ along $\langle 1, \pm 1, 1\rangle$. Each manifold corresponds to a value of σ_θ:

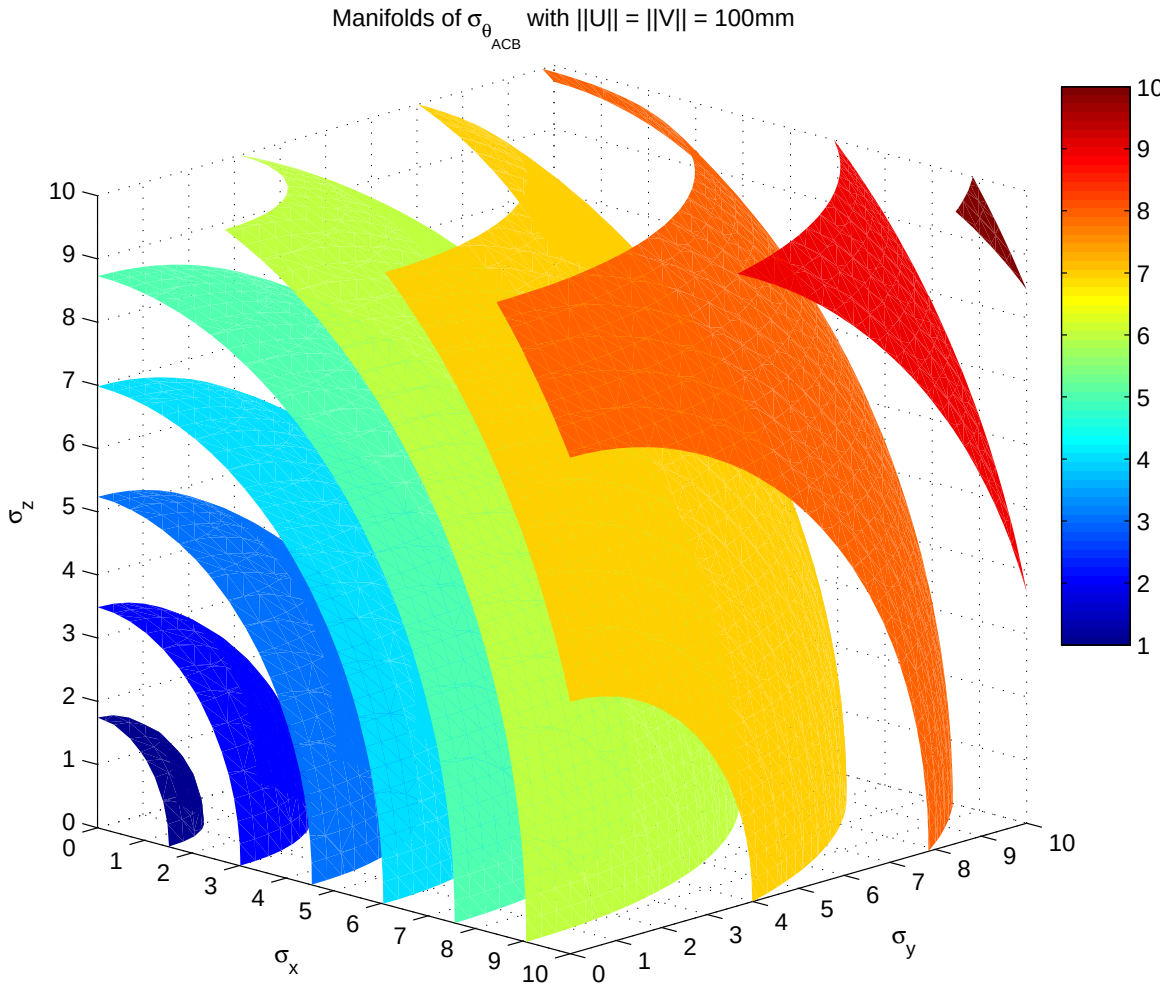


Figure 4.9: Distinct manifolds of σ_{ACB} for varying marker axes errors (axes are scaled in degrees)

Lastly, given the nature of the covariance matrix, $\mathbf{\Lambda}_{\vec{X}}$, it can be separated so that the individual effects on the angular deviation can be calculated:

$$\begin{aligned} \mathbf{\Lambda}_{\vec{X}} &= \mathbf{\Lambda}_{\vec{X}}|_A + \mathbf{\Lambda}_{\vec{X}}|_B + \mathbf{\Lambda}_{\vec{X}}|_C \\ &= \begin{bmatrix} \mathbf{\Lambda}_{world} & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset \end{bmatrix} + \begin{bmatrix} \emptyset & \emptyset & \emptyset \\ \emptyset & \mathbf{\Lambda}_{world} & \emptyset \\ \emptyset & \emptyset & \emptyset \end{bmatrix} + \begin{bmatrix} \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \mathbf{\Lambda}_{world} \end{bmatrix} \end{aligned} \tag{4.25}$$

Thus, the contributions of each of the three points to the angular error can be calculated. It is further possible to separate $\mathbf{\Lambda}_{\vec{X}}$ such that the effects per axis or per plane can be calculated.

This separation was not done as part of this thesis, as it would be used for further analysis of the system that is not part of the scope of this thesis; it is mentioned because it results from the mathematics proven herein.

Chapter 5

The Physiotherapy Study and Use of the PMM

The focus of this thesis is on the workings of the PMM: its stereo-vision theory and the underlying mathematics that make the system run. The focus is not geared towards analyzing the actual application of the PMM beyond meeting the specifications required by the physiotherapists. Yet, while the application was not a primary focus, it did have a large influence on the design considerations. The main application of the PMM in the Department of Mechanical and Mechatronic Engineering is its future use as an experimental platform to gauge how different methods and algorithms in computer vision affect the system. However, for the Division of Physiotherapy, the PMM has to be a tangible tool for the accurate measurement of patient posture; each of these aspects mould the system.

An important consideration is that, during the design and improvement process, the initial PMM was already put in the field to collect the initial data for the physiotherapy study of Yolandi Brink [31].

5.1 Initial Phases – PMM in Design

The PMM as experimental platform would simply be the cameras connected and configured in a laboratory space – to be reconfigured as required to meet the specific purpose of what ever experiment is to be run on it; in the laboratory the system mutable. But, as a portable system that tours schools, the PMM would need to be tailored to meet specific needs.

To simply say that the machine must be portable is an understatement; in truth, it has to be carried from a place of storage, transported in a car, carried to the relevant classroom through a school building – navigating corridors, stairwells and doorways – before being set up in an environment that changes from school to school.

These considerations had to be contemplated before the machine could begin to travel to schools. Meetings with the physiotherapists were important; they had previously been to the schools for subject screening and were familiar with the spatial dynamics they wanted investigated in terms of how the subjects would be seated during capture.

In order to assist Brink with the PMM set-up and usage, as well as to use the opportunity to refine the design of the PMM, the author agreed to volunteer as the technical assistant in Brink's study.

5.2 Data Capturing – PMM in Use

From an engineering perspective, the capture volume contains everything inside it, whether markered for capture or not, but for the physiotherapists the only data desired is that of the markers. Thus the two ways of thinking of the capture volume are important so that crucial aspects are not overlooked, i.e., different sized desks obscuring the cameras' view of certain markers, or the capture volume being unevenly illuminated. The effect of these aspects became apparent during the use of the system at different locations in different environments.

Between June 2009 and June 2010, the PMM visited eighteen schools on twenty-six separate occasions, capturing approximately 400 sets of data (including calibration sets and follow-up captures of the same learner). As compressed JPG images, this amounts to some 7 GBytes of unprocessed data. These visits as part of the physiotherapy study formed a tour of the highschools of the Greater Cape Town area; from poor, ill-equipped schools to affluent, well-equipped schools – presenting a varying range of capture environments.

Figure 5.1 shows four examples of school classrooms encountered during the study.

(a) An adequately spaced and lit classroom.

(b) Another adequately spaced and lit classroom

(c) A cramped classroom

(d) Excessive light reflections off background walls

Figure 5.1: examples of school classrooms used during the physiotherapy study of Brink [31]. The images were taken as part of the calibration and were chosen to illustrate capture volumes that can typically be encountered.

The PMM's current form is largely due to the author being able to assess its useability in the different school environments, including all aspects from transporting it to and from the schools, to setting it up inside the classroom.

Due to the nature of medical studies, any further suggested improvements to the PMM cannot be implemented until the follow-up study has been completed, as the machine may not change form lest it affect the results of the study. All further suggestions to form and function are discussed in §7.

Chapter 6

The *Mannequin Test* Experiment to Confirm Accuracy

Having shown that the PMM is able to capture applicable images for the physiotherapy requirements, and having shown that the suggested improvements to the PMM are tangible improvements, the important step now is to show the accuracy of the machine itself.

Confirming the accuracy is not only important from an engineering perspective, but also for the physiotherapists – in fact more so for the physiotherapists. In engineering, a lot of mathematics is commonplace and many concepts seem to be self-evident; however, in the medical field there are deeper, more rigourous ethical considerations when using a machine. For example, if a machine is tested with a specific age group, then the use of that machine may only be used for subjects of that age group. For the machine to be used with subjects from a different age group it must first be *validated* with the different age group [32]. Proper validation assures the accuracy of the machine for its intended purposes by producing tangible evidence to illustrate the accuracy.

In order to have formal proof of the accuracy of the PMM, the *mannequin test* was created in collaboration with Yolandi Brink of Physiotherapy Division.

6.1 Premise and Experimental Setup

The PMM needs to capture a set of accurately known points and then calculate the positions of these known points. Concurrently, the VICON Imaging System (available at the Physiotherapy Division) would also capture this set of points; the

difference between the two data sets would show how accurate the PMM is when compared to the VICON.

The use of the VICON is widespread — it used in many research papers as a golden standard for motion analysis. However, the VICON company does not publish direct data on the accuracy of their systems, mainly because it can vary greatly from laboratory to laboratory and set-up to set-up. As a result, the VICON is compared to a Computer Measuring Machine (CMM), through the use of the pyramid calibration object, in §6.2.

Once the accuracy of the VICON is established, the PMM can be compared to the VICON – as done in §6.3. Further, the accuracy of the PMM compared to CMM is possible through the use of the Calibration Object and back-projection, as discussed in §6.4.

6.1.1 Mathematical Aspects

A set of real world points, $S = \left\{\vec{X}_0, \ldots, \vec{X}_n\right\}$, must be captured by the PMM and triangulated to get $S|_{PMM}$. Independently, these points in S should be accurately calculated by another, trusted mechanism, e.g., VICON, Computer Measure Machine (CMM), etc., to get $S|_{compare}$.
A real world point is defined as:

$$\vec{X} = \langle X, Y, Z \rangle \tag{6.1}$$

An image point is defined as:

$$\vec{x} = \langle x, y \rangle \tag{6.2}$$

The absolute system errors (ε) can then be expressed in all three spatial dimensions, and an absolute sum error can be used to compare the accuracy between different sets of captured data.

$$\begin{aligned} \varepsilon_{sets} &= \|S|_{PMM} - S|compare\| \\ &= \left\{\left\|\vec{X}_{1_{PMM}} - \vec{X}_{1_{compare}}\right\|, \ldots, \left\|\vec{X}_{n_{PMM}} - \vec{X}_{n_{compare}}\right\|\right\} \end{aligned} \tag{6.3}$$

Thus, per axis, the absolute errors are:

$$\varepsilon_{\hat{i}} = \left\{\left|x_{1_{PMM}} - x_{1_{compare}}\right|, \ldots, \left|x_{n_{PMM}} - x_{n_{compare}}\right|\right\} \tag{6.4}$$

$$\varepsilon_{\hat{j}} = \left\{\left|y_{1_{PMM}} - y_{1_{compare}}\right|, \ldots, \left|y_{n_{PMM}} - y_{n_{compare}}\right|\right\} \tag{6.5}$$

$$\varepsilon_{\hat{k}} = \left\{\left|z_{1_{PMM}} - z_{1_{compare}}\right|, \ldots, \left|z_{n_{PMM}} - z_{n_{compare}}\right|\right\} \tag{6.6}$$

The sum of the absolute error, used as a metric to quantify the total error in a system or subsystem, is taken as:

$$SAE_{set} = \sum_{i=1}^{n} \varepsilon_i,\ \varepsilon_i \in \varepsilon_{set} \tag{6.7}$$

The mean value of a set is take as the arithmetic mean, μ_{set}, and can also be written as $\bar{x} = \mu_x$.

$$\mu_{set} = \frac{1}{n} \sum_{i=1}^{n} s_i,\ s_i \in S_{set} \tag{6.8}$$

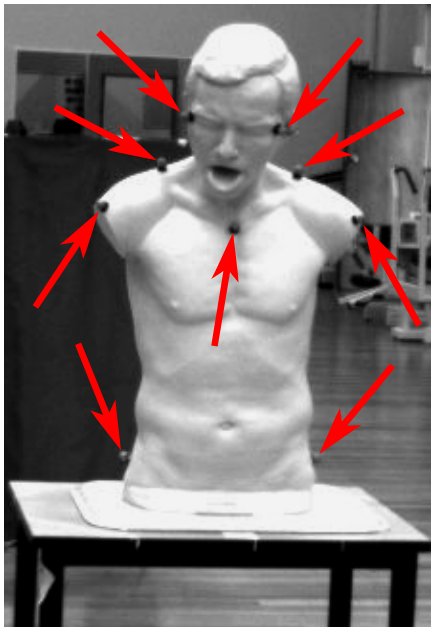
Where the sets are taken as Gaussian distributions, the standard deviation, σ_{set}, is taken as:

$$\sigma_{set} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (s_i - \mu_{set})^2} \tag{6.9}$$

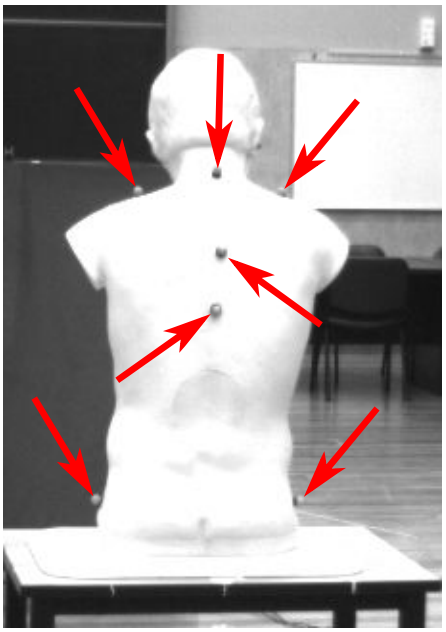
The value, σ_{set}^2, is called the *variance* of the set.

6.1.2 Set-up

The PMM was set up in the Physiotherapy laboratory on the Tygerberg Health Sciences Campus in order to make use of the VICON. The capture subject used for this experiment was a *"Choking Charlie" Heimlich Abdominal Thrust Maneuver Training Mannequin*, seen markered in Figure 6.1.



Model Front



Model Back

Figure 6.1: Markered mannequin model ready for capture (markers indicated by the red arrows)

The VICON and PMM are not synchronized together, and as a result they must be triggered independently by two operators. Because of this, capture frames from the PMM must be matched to the VICON. It was felt that a fixed object would be a better capture subject than a markered human, as this would reduced any possible movement of the subject and thus reduce the need to match capture frames.

The idea of using a mannequin had been discussed previously with Brink; the idea was revived, as a mannequin is a static, rigid object in the shape of a human. This means that it would be a suitable representation of the applicable capture volume for the PMM, and that marker placement would be familiar to Brink. Mathematically, any rigid object would have been acceptable, but the added benefit of the mannequin was that it aided the design process.

6.1.3 Collecting the Data

The PMM was calibrated with the pyramid ground-truth object. The VICON was calibrated using the VICON 5-point L-wand. The VICON cameras were T120s with a Giganet controller. The VICON software was Nexus (version 1.5.2).

The pyramid object was also captured (once) by the VICON, in this way the accuracy of the VICON could be compared directly to the CMM through world points of the calibration object.

The mannequin was captured nine times. For each successive capture the mannequin was placed in a different world orientation, either tilted, rotated or a combination of each. Once again, the idea was to see how differing postures would shift about the capture volume. This further helped in the evaluation of the best orientation of the cameras around the markered human form in order to create the capture volume.

Once the data had been captured, it needed to be processed. The relevant VICON data was labelled and exported. Brink used the PMM images to perform marker selection for calibration and reconstruction. Further, the PMM images were digitized independently of Brink while testing the software design; this made it possible to test between two separate users digitizing the images. No significant difference was found between the two independent digitizations, when compared to two separate digitizations of the same data by the same user; in some cases the differences could be as less than a millimetre to as much as ten millimetres. The user selection error needs to be studied further to better quantify how repeatable human marker selection is.

6.1.4 Transformation between Reference Frames

The PMM data and the VICON data have two different coordinate systems as a result of their independent calibration. As a result, the VICON coordinate frame must be transformed to the frame of the calibration object. Thus a coordinate frame transformation matrix, $\boldsymbol{T}$, must be found such that:

$$T_{E/O} \quad : \quad {}^{E}\vec{X}_{\nu_i} \mapsto {}^{O}\vec{X}_{\nu_i} \tag{6.10}$$

The reverse map also holds; it is the inverse of the forward map and also the transpose of the forward map; the coordinate frame transformation matrix is symmetrical:

$$T_{O/E} \quad : \quad {}^{O}\vec{X}_{\nu_i} \mapsto {}^{E}\vec{X}_{\nu_i} \tag{6.11}$$

$$= \quad (T_{E/O})^{-1} = (T_{E/O})^{\mathsf{T}} \tag{6.12}$$

If there is a translation $\vec{T}_r$ between the two reference frames, then this must be accounted for either before or after the transformation – depending on the transformation direction. The mathematical explanation can be found in Appendix B.

6.2 Confirming the Accuracy of the VICON

During the calibration stage of the PMM, the VICON was used to capture the pyramid calibration object, as discussed in §6.1. Two data sets are available from this: the VICON capture of the calibration object, relative to the origin of the VICON's calibration,

$$^{E}D_{\nu} = \left\{ \vec{X}_{\nu_1}, \ldots, \vec{X}_{\nu_n} \right\},$$

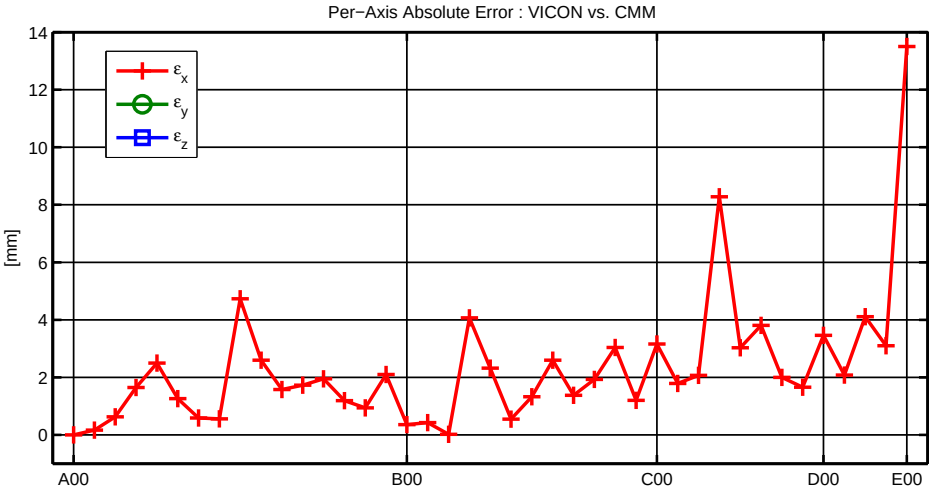
and the CMM ground truth definition data set from when the object was measured on the computer measuring machine,

$$^{O}D_{\Omega} = \left\{ \vec{X}_{\Omega_1}, \ldots, \vec{X}_{\Omega_n} \right\}.$$

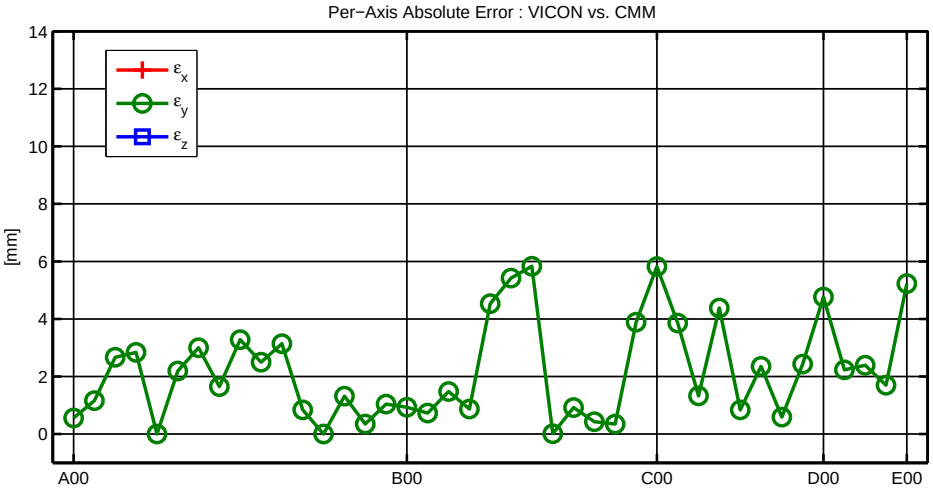
These two sets were used to find the transformation from VICON frame to PMM frame; all VICON data points were transformed to the O frame:

$$^{E}D_{\nu} \mapsto {}^{O}D_{\nu}$$

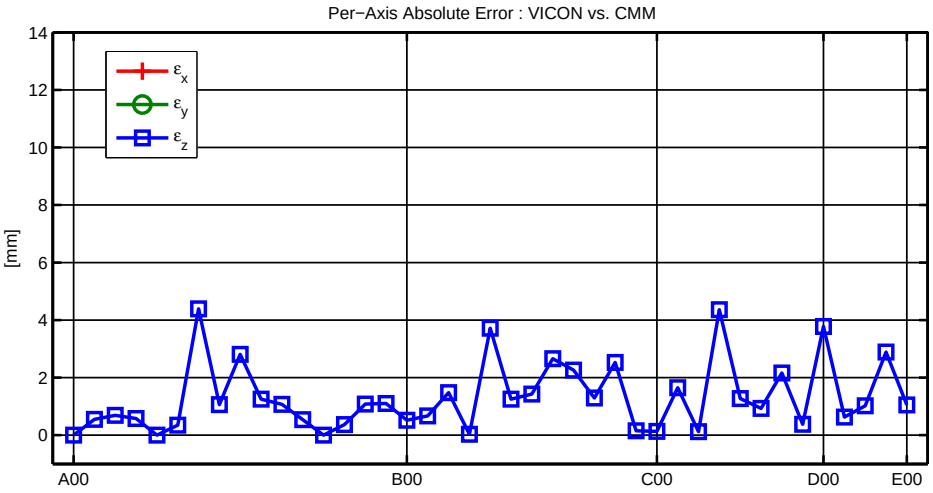
Thereafter, for each point on the calibration object, the error between the VICON reported value and the CMM measured value was calculated. Figure 6.2 shows the absolute errors per axis for each landmark on the calibration object:



(a) Error along X-axis



(b) Error along Y-axis



(c) Error along Z-axis

Figure 6.2: Absolute errors per axis for the VICON vs. CMM

As the pyramid calibration object is a tiered object, in that it has five distinct layers, $A00$-$A15$, $B00$-$B11$, $C00$-$C06$, $D00$-$D03$ and $E00$, that differ in height (A lowest, E highest), these can be taken as distinct subsets. During measuring on the CMM it was felt that, due to the length of the dowels in the pyramid, the taller tiers may move during CMM measurement and thus could be unreliable. The error for $E00$, as seen in Figure 6.2, is substantial in comparison to the other values and justifies the decision not to use the point in the comparison of the VICON and the CMM.

Table 6.1 shows the statistics for the absolute error between the VICON and the CMM for the entire pyramid calibration object, levels A to E.

Table 6.1: Absolute error statistics for the VICON vs. CMM : $\boldsymbol{A00 - E00}$

$\boldsymbol{A00–E00}$	ε_x	ε_y	ε_z
μ	2.33 mm	2.19 mm	1.32 mm
σ	2.36 mm	1.71 mm	1.21 mm
SAE	95.50 mm	89.83 mm	54.23 mm
# Markers	41	41	41

Excluding level E (which only has one point, $E00$), the comparison shows a marked difference in the x-axis standard deviation, as seen in Table 6.2:

Table 6.2: Absolute error statistics for the VICON vs. CMM : $\boldsymbol{A00 - D03}$

$\boldsymbol{A00–D03}$	ε_x	ε_y	ε_z
μ	2.05 mm	2.12 mm	1.33 mm
σ	1.57 mm	1.66 mm	1.22 mm
SAE	82.00 mm	84.60 mm	53.19 mm
# Markers	40	40	40

Due to the uncertainty about the dowels moving during the CMM measurement, it is important to look at the errors of the dowels that did not move visibly, namely tiers A and B, as show in Table 6.3.

Table 6.3: Absolute error statistics for the VICON vs. CM : $\boldsymbol{A00 - B11}$

$\boldsymbol{A00–B11}$	ε_x	ε_y	ε_z
μ	1.55 mm	1.85 mm	1.21 mm
σ	1.17 mm	1.63 mm	1.14 mm
SAE	43.45 mm	51.93 mm	33.86 mm
# Markers	28	28	28

This shows a marked difference — an improvement in the accuracy of the comparison, especially in the x-axis.

Given that the VICON cameras are situated between 3.5 m to 4 m away from the calibration volume, and that the markers on the pyramid object are of a cruder construction than would be preferable for detection by the VICON, it is deemed that a 1 mm to 2 mm average error would be acceptable. However, further attempts should be undertaken to accurately quantify the VICON error and improve the pyramid calibration object.

6.3 Comparing the PMM with the VICON

With calibration complete, the mannequin was placed in different orientations within the capture volume. A total of nine sets were captured, orientated as in Table 6.4:

Table 6.4: Names and descriptions of data sets

Set	Set Name	Description
P001	CFacingY	Facing forward along positive y-axis.
P002	CFacingX	Facing forward along positive x-axis.
P003[1]	CFacingMY	Facing forward only negative y-axis.
P004	CRot135CCW	Rotated 225° from x-axis.
P005	CFacingYTiltF	Facing forward along positive-axis; tilted forward.
P006	CFacingYTiltFL	Facing forward along positive-axis; tilted left.
P007	CRot45CWTiltFL	Rotated 45° from x-axis; tilted left.
P008	CFacingYRemarker01	Facing forward along positive y-axis.
P009	CFacingYRemarker02	Facing forward along positive y-axis.

For each capture, a VICON data set and a PMM data set were created. The VICON data set was transformed to the PMM reference frame. Note that not all sets have an equal number of markers; this is due to markers either falling off during a capture, not being seen by sufficient VICON cameras for a reconstruction, or nothing being seen by sufficient VICON cameras for a reconstruction.[2]

Table 6.5 shows the statistical data, per capture set and per axis, for the PMM compared to the VICON:

[1]Note that the data for *P*003 was corrupt and could not be used.

[2]On the day of this experiment, one of the VICON cameras was out-of-order, this limited the VICON capture volume.

Table 6.5: Per-set statistical data for the PMM compared to VICON

Average error per set per axis [mm]								
	P001	**P002**	**P004**	**P005**	**P006**	**P007**	**P008**	**P009**
$\bar{\varepsilon}_x$	8.04	3.67	3.94	7.13	6.42	8.79	4.08	5.61
$\bar{\varepsilon}_y$	17.62	0.72	1.15	1.01	0.72	2.57	0.74	0.82
$\bar{\varepsilon}_z$	2.69	2.03	2.48	2.33	2.16	3.74	2.73	2.88

Standard deviation per set per axis [mm]								
	P001	**P002**	**P004**	**P005**	**P006**	**P007**	**P008**	**P009**
$\boldsymbol{\sigma}_x$	7.96	1.65	3.47	7.15	5.45	6.06	4.13	3.90
$\boldsymbol{\sigma}_y$	27.93	0.38	0.91	0.90	0.51	3.42	0.68	0.59
$\boldsymbol{\sigma}_z$	1.00	1.11	0.94	1.54	1.39	1.91	1.41	1.59

Sum of the absolute errors per set per axis [mm]								
	P001	**P002**	**P004**	**P005**	**P006**	**P007**	**P008**	**P009**
$\boldsymbol{SAE}_x$	112.56	44.08	55.10	99.82	89.92	114.32	53.00	78.56
$\boldsymbol{SAE}_y$	246.67	8.68	16.15	14.20	10.02	33.44	9.67	11.41
$\boldsymbol{SAE}_z$	37.67	24.31	34.71	32.60	30.21	48.60	35.52	40.36

Number of markers in each set [ ]								
	P001	**P002**	**P004**	**P005**	**P006**	**P007**	**P008**	**P009**
$\#_x$	14	12	14	14	14	13	13	14
$\#_y$	14	12	14	14	14	13	13	14
$\#_z$	14	12	14	14	14	13	13	14

As can be seen, the data is varied, but not drastically so – except for $P001$ all other values are below 10 mm, and the sum of the absolute errors, the SAE values, indicate that the sets are comparable and do not differ greatly from each other.

Due to the fact that $P001$ in the y-axis is an order of magnitude higher than the rest of the sets, as highlighted in Table 6.5, $P001$ will be marked as an outlier.

The total statistical data for all the sets together was calculated as one large set. This gives the overall accuracy of the machine across successive captures. Table 6.6 shows the per-axis statistics for all the captured data, including the outlier set, $P001$:

Table 6.6: Total statistical data for all captures as one set:
$Data = \{P001, P002, P004, P005, P006, P007, P008, P009\}$

Total	ε_x	ε_y	ε_z
$\boldsymbol{\mu}_{Total}$	5.99 mm	3.24 mm	2.63 mm
$\boldsymbol{\sigma}_{Total}$	5.52 mm	11.30 mm	1.44 mm
$\boldsymbol{SAE}_{Total}$	647.35 mm	350.24 mm	283.98 mm
$\#_{Total}$	108	108	108

The total statistical data can then be compared with the statistical data excluding the outlier, $P001$, as shown in Table 6.7:

Table 6.7: Total statistical data excluding set $P001$:
$Data = \{P002, P004, P005, P006, P007, P008, P009\}$

Total	ε_x	ε_y	ε_z
$\boldsymbol{\mu}_{Total}$	5.69 mm	1.10 mm	2.62 mm
$\boldsymbol{\sigma}_{Total}$	5.05 mm	1.51 mm	1.49 mm
$\boldsymbol{SAE}_{Total}$	534.79 mm	103.57 mm	246.31 mm
$\#_{Total}$	94	94	94

As can be seen, there is a substantial reduction in the mean error (66%) and standard deviation (87%) of the mean error along the y-axis; the improvement to the standard deviations is imperative. As discussed in §4 on error propagation, it is the variance that propagates through to the angular error variance. The mean error certainly plays a role, but it is the three-sigma, 3σ, that ultimately dominates. Statistically, 99.7 % of all recorded values fall within the 3σ boundary; 95 % within 2σ and 68 % within 1σ. Thus the larger the standard deviation, the more spread out the points will be.

Furthermore, the shortest distance specified for the physiotherapy study is the canthus to tragus distance (eye to ear) – in the case of the author this is approximately 76 mm. With the statistics given in Table 6.7 and the error model of §4, the expected error for an angle between two vectors, both of length 75 mm, is $\varepsilon_\theta = 3.07°$, which is smaller than the 5° limit that forms part of the physiotherapy specification. The smallest vector length to achieve a 5° error angle is 46 mm.

The error ellipsoids are used to illustrate the importance of the standard deviation (as well as the mean error). Given a point $\vec{x}_i$, an uncertainty boundary can be drawn around the point as defined by the mean error and standard deviation of that axis; this boundary would represent the region that the point $\vec{x}_i$ is most likely to be in. The boundary is an ellipsoid with elliptical axes given by $\bar{\varepsilon}_{axis} + k \cdot \sigma_{axis}$, where $k \in [0, \infty)$ and corresponds to the probability of being in that boundary, $k = 1 \rightarrow 68\,\%$ etc., as previously discussed. The concept of the uncertainty boundary is shown in Figure 6.3:

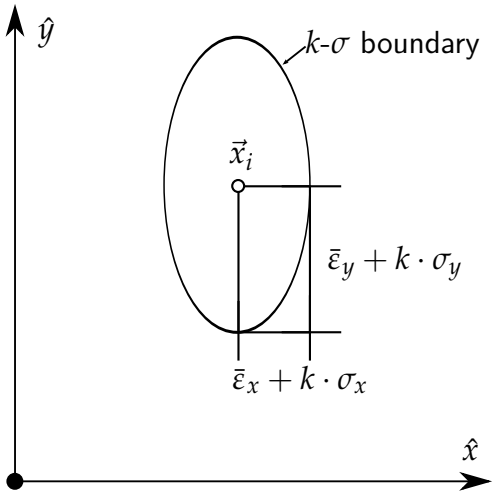


Figure 6.3: Definition of an error ellipsoid

To illustrate the accuracy of the PMM compared to the VICON, Figure 6.4 shows the uncertainty boundaries for the statistical data in Table 6.7, overlaid onto the 15 mm diameter marker (grey dashed circle) in the three 3D planes: xy, xz, yz. The scale is 1 : 1 in millimetres and the outermost ellipse corresponds to the 3σ–99.7 % boundary, going inwards to 0σ, which is the mean error.

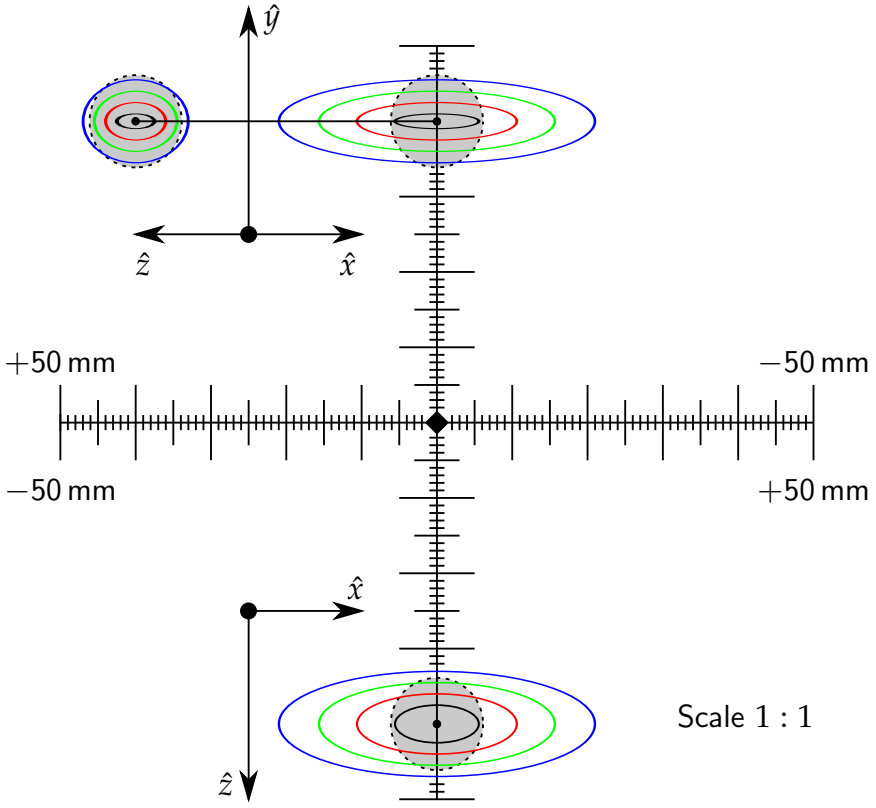


Figure 6.4: Error ellipsoid for the PMM vs. VICON data superimposed on top of the 15 mm diameter markers in the xyz planes

This data clearly shows that the x-axis has the largest associated error and statistical deviation. This is due to camera orientations around the origin, the focal axis of the cameras which are closest aligned to the x-axis of the capture volume with have the larger associated x-axis error. It is an expected effect that corroborates well with established theory [15], as explained in §3.3.2.

This comparison of the PMM to the VICON shows that the PMM is capable of achieving the desired accuracy under similar conditions to those that it is expected to operate in. Generally, the PMM cameras are between 1.5 m and 2.5 m away from the capture subject and are spaced approximately 0.5 m apart on their rigid bar. These results are favourable and show promise for further development and improvement of the PMM.

6.4 Accuracy of the PMM through Back-Projection

With the PMM compared to the VICON and the VICON compared to the CMM, there is still an associated uncertainty on the PMM. Is the uncertainty of the VICON with regard to the CMM affecting the uncertainty of the PMM to the VICON favourably or not?

Ideally, the PMM should be compared directly to the CMM through the use of another object similar to the calibration object – this would accurately determine the uncertainties of the system. The drawback, however, is that a second calibration object would need to be carted about as part of the system; after all, these uncertainties per capture volume would be used to predict angular uncertainties of the PMM.

Fortunately there is the calibration object – already measured on the CMM. It could be argued that, because the system is calibrated with the calibration object, using the same object would yield the initial points. Mathematically, this would be true for ideal points with an ideal camera matrix. However, real world applications are not ideal and thus yield the errors that are used to determine the uncertainty of the system. Further, it must be noted that the computational algorithms (e.g. SVD) are approximations to the ideal homographies and inverse matrices, and as a result have an inherent approximation error. Due to these factors, the camera matrix will differ from the ideal; this error determines the uncertainty of the system.

6.4.1 Image Pixel Error

All of these errors cause the calibrated camera matrix to differ slightly from the ideal; thus the CMM-measured ground-truth points can be back-projected onto the image planes and then compared to the initial image points selected for calibration – as implemented in pyMultiCam [25]. The statistical results of this back-projection can be seen in Table 6.8, per camera, per image axis, in pixels.

Table 6.8: Per-camera back-projection pixel error

Average error per camera per image axis [px]					
	C0	**C1**	**C2**	**C3**	**C4**
μ_x	0.15	0.38	0.23	0.25	0.34
μ_y	0.22	0.61	0.29	0.31	0.30

Standard deviation per camera per image axis [px]					
	C0	**C1**	**C2**	**C3**	**C4**
σ_x	0.09	0.28	0.16	0.17	0.26
σ_y	0.15	0.44	0.22	0.25	0.18

Sum of absolute errors per camera per image axis [px]					
	C0	**C1**	**C2**	**C3**	**C4**
SAE_x	3.66	9.53	5.64	6.15	8.51
SAE_y	5.50	15.36	7.26	7.87	7.54

Number of points in each set [ ]					
	C0	**C1**	**C2**	**C3**	**C4**
$\#_x$	25	25	25	25	25
$\#_y$	25	25	25	25	25

The overall statistical results for the PMM are presented in Table 6.9, per axis, in pixels.

Table 6.9: Total statistical data for back-projection error

Total	ε_x	ε_y
$\boldsymbol{\mu_{Total}}$	0.27 px	0.35 px
$\boldsymbol{\sigma_{Total}}$	0.22 px	0.30 px
$\boldsymbol{SAE_{Total}}$	33.49 px	43.52 px
$\boldsymbol{\#_{Total}}$	125	125

To better interpret this data, Figure 6.5 illustrates the error ellipses for the five cameras. Note that the centre points are arbitrarily assigned; the centres correspond to the top-left point of the pixel – pixels being the entire region between grid-intersections in the figure. Thus the ellipses show full sub-pixel error distribution. The innermost ellipse corresponds to the 50% confidence boundary, the others increase outwards from 1σ to 3σ.

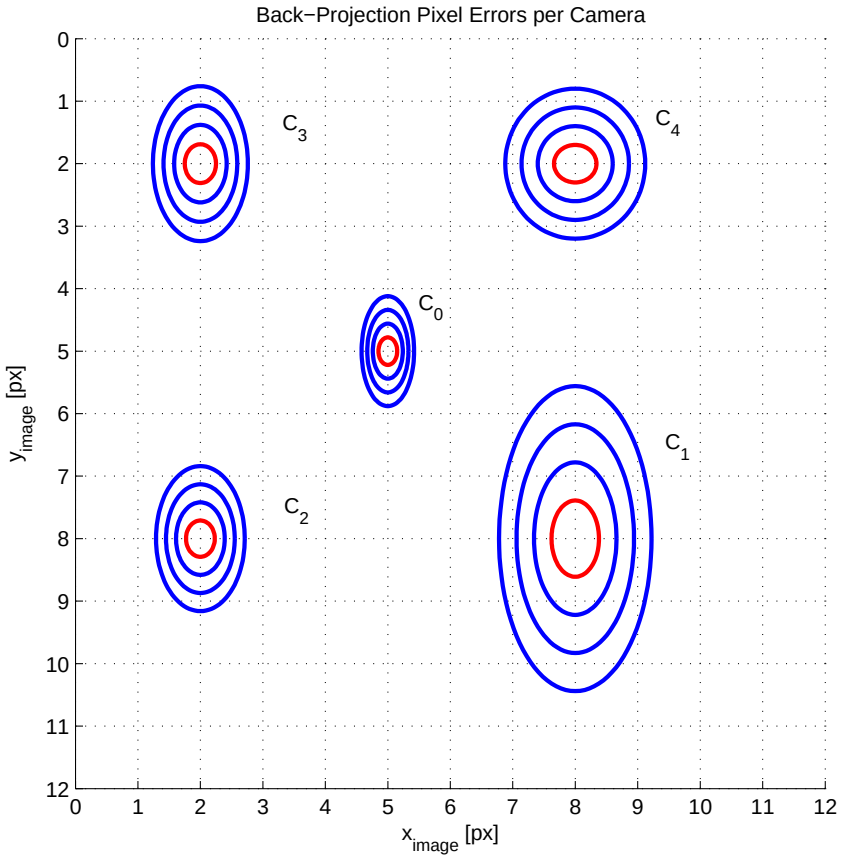


Figure 6.5: Pixel error ellipses per camera

Bear in mind that the region in the figure is a small section of an image that is in fact 640×480 px in dimension; this means that there is very little error in the whole image. However, the capture volume does not represent the entire image and these errors only apply to the region of the image plane that would correspond to the capture volume projection onto that image plane.

The importance of these image pixel errors is that they illustrate which cameras in the system may be problematic; it can be seen in Figure 6.5 that camera C_1 is comparatively the worst camera in the system, and therefore it can be expected that triangulation pairs with C_1 in would have the larger errors.

6.4.2 Camera Pair Calibration Error

In the same way that that the CMM points can be back-projected, so can the image points for calibration be triangulated forward and compared with the CMM ground-truth points. This would give the real world uncertainties per camera pair and can be used to determine the overall real world uncertainties of the system.

Table 6.10 represents the statistical data for the stereo pairs available in the system — calculated for all levels in the calibration object, A–E.

Table 6.10: Per-camera pair error statistics (*A*–*E*), from camera calibrations and triangulation

Average error per camera pair per axis [mm]				
	C0:C1	**C0:C2**	**C1:C2**	**C3:C4**
$\bar{\varepsilon}_x$	6.74	1.77	4.80	3.32
$\bar{\varepsilon}_y$	1.49	0.43	0.90	0.66
$\bar{\varepsilon}_z$	1.50	0.73	1.09	0.66

Standard deviation per camera pair per axis [mm]				
	C0:C1	**C0:C2**	**C1:C2**	**C3:C4**
$\boldsymbol{\sigma}_x$	5.20	0.98	3.33	2.77
$\boldsymbol{\sigma}_y$	1.03	0.36	0.82	0.47
$\boldsymbol{\sigma}_z$	1.17	0.48	0.85	0.63

Sum of absolute errors per camera pair per axis [mm]				
	C0:C1	**C0:C2**	**C1:C2**	**C3:C4**
$\boldsymbol{SAE}_x$	168.62	44.25	119.94	83.00
$\boldsymbol{SAE}_y$	37.16	10.87	22.46	16.49
$\boldsymbol{SAE}_z$	37.43	18.17	27.30	16.43

Number of points in each set [ ]				
	C0:C1	**C0:C2**	**C1:C2**	**C3:C4**
$\#_x$	25	25	25	25
$\#_y$	25	25	25	25
$\#_z$	25	25	25	25

The x-axis consistently is the worst performing axis; once again this is to be expected. Note that the pairs containing camera C_1 have higher mean errors and higher standard deviations, which are consistent when considering the pixel errors for camera C_1.

The total statistical data for the entire system is presented in Table 6.11:

Table 6.11: Total statistical data of expected errors from the system (*A*–*E*) as calculated from camera calibrations and triangulation

Total	ε_x	ε_y	ε_z
$\boldsymbol{\mu}_{Total}$	4.16 mm	0.87 mm	0.99 mm
$\boldsymbol{\sigma}_{Total}$	3.88 mm	0.82 mm	0.89 mm
$\boldsymbol{SAE}_{Total}$	415.81 mm	86.97 mm	99.33 mm
$\#_{Total}$	100	100	100

Comparing these values of PMM to PMM (through the calibration object) in Table 6.11 with that of PMM to VICON in Table 6.6, it shows that the PMM is more accurate than expected in comparison with the VICON.

For consistency, the *E* level was removed from the calibration image points and the system was re-calibrated. The individual point data was not reprocessed; however, Table 6.12 shows the statistical data for the total system as determined by triangulation compared to the ground-truth points.

The percentage improvement indicated compares the data in Table 6.12 to Table 6.11 (the data for PMM to PMM including *E*) – a positive value is a decrease in error and a negative value is an increase in error:

Table 6.12: Total statistical data of expected errors from the system (*A*–*D*) as calculated from camera calibrations and triangulation

Total	ε_x	ε_y	ε_z
$\boldsymbol{\mu}_{Total}$	3.02 mm	0.64 mm	0.76 mm
$\boldsymbol{\sigma}_{Total}$	2.65 mm	0.55 mm	0.56 mm
$\boldsymbol{SAE}_{Total}$	290.25 mm	61.21 mm	72.75 mm
$\#_{Total}$	96	96	96
$\boldsymbol{\mu}_{\%}$	27.29 %	26.68 %	23.71 %
$\boldsymbol{\sigma}_{\%}$	31.70 %	33.45 %	37.13 %

There was a decrease of roughly a 30 % in errors across all axes; the largest improvement in mean error was in the x-axis and the largest improvement in standard deviation was in the z-axis. The x-axis still remains the axis with the greatest error.

Previously, when the error ellipses were plotted they were defined as if the standard deviations of each axes were independent of each other; however, this is not entirely true. In the case of the camera's 3D errors: if the focal axis is not perfectly parallel to one of the world axes, then errors along the focal axis will result in errors along all three world axes, where the variance σ_x^2, σ_y^2, σ_z^2 describes how far the errors are spread from the mean error. The covariance, σ_{xy}^2, then describes how changes in x influence y and vice versa [33]. $\mathbf{\Delta}_{set}$ is the covariance matrix, and is used to determine the angular uncertainty, as discussed in §4. It can also be used to create a more accurate representation of the error ellipsoids. The covariance modifies the ellipsoids slightly to represent how changes in the one axis influence the other axes. Figures 6.6 to 6.9 show error ellipsoids for the four camera pairs in XY, XZ, YZ and $3D$ respectively.

The effects of the covariance σ_{xy}^2 on the camera pair C_0C_1 can be seen clearly in Figure 6.6:

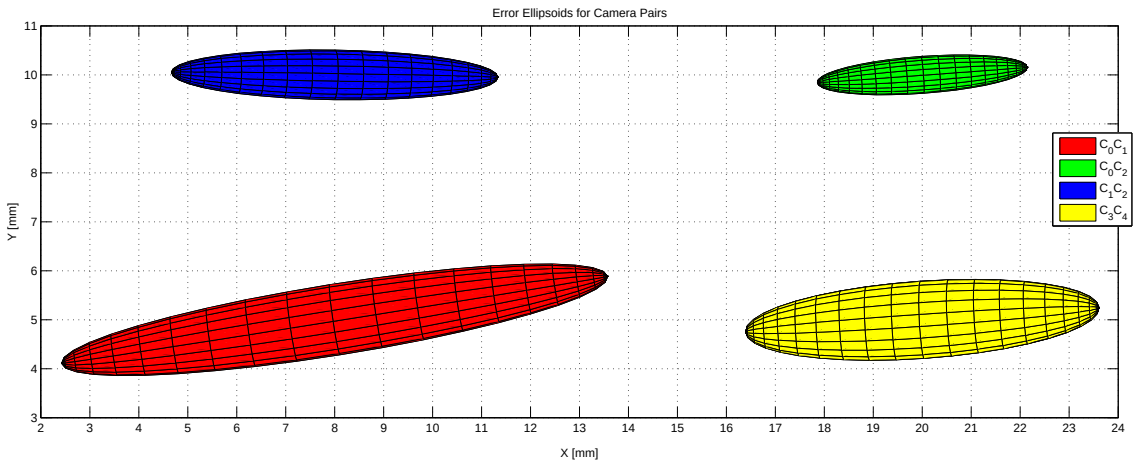


Figure 6.6: Error ellipses per camera pair as calculated from the covariance matrix – XY view

The σ_{xz} covariance is slight in camera pair C_0C_2:

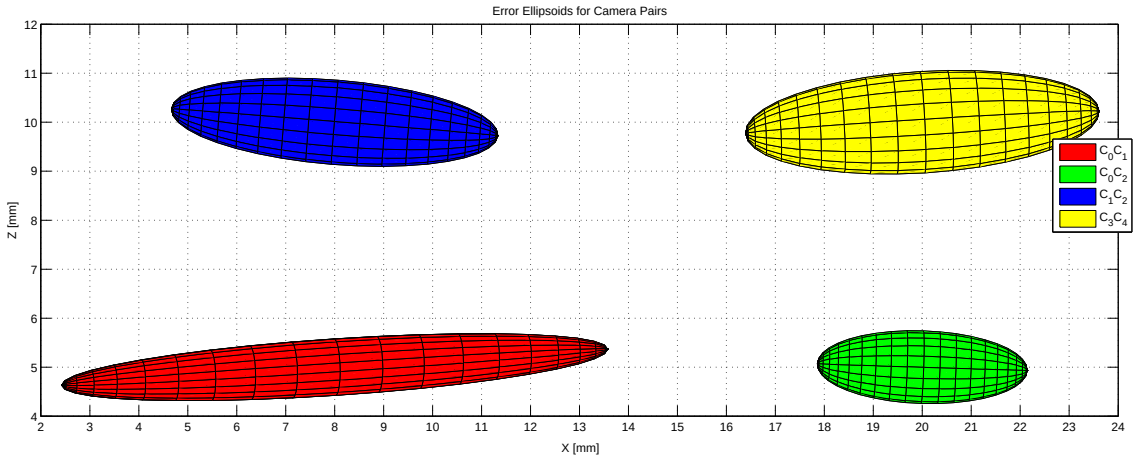


Figure 6.7: Error ellipses per camera pair as calculated from the covariance matrix – XZ view

The σ_{yz} covariance is pronounced in camera pair C_0C_1:

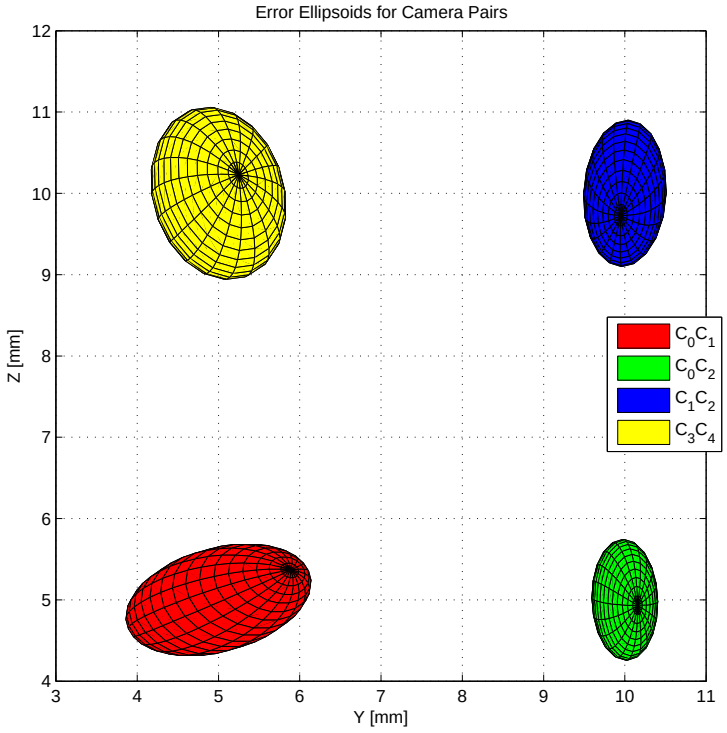


Figure 6.8: Error ellipses per camera pair as calculated from the covariance matrix – YZ view

The 3D view gives a spatial perspective to these camera pair error distributions. The plots are in millimetres and show a very tight distribution in the y and z axes; the x-axis still remains the axis with the largest error. However, due to the covariances it cannot be assumed that the errors in the three axes are independent.

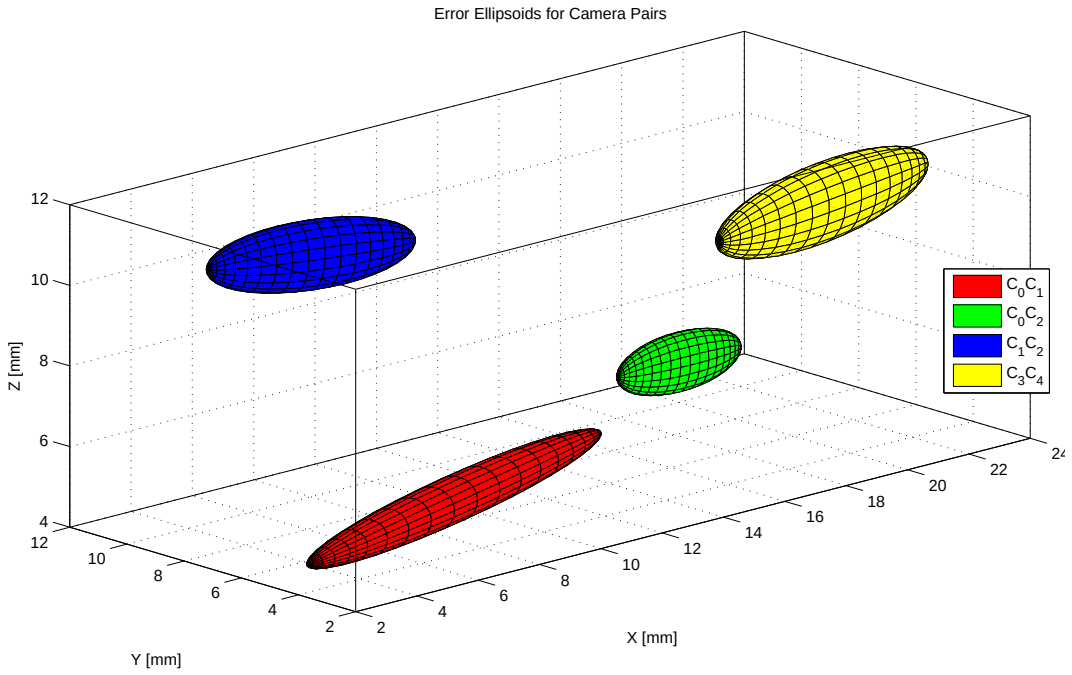


Figure 6.9: Error ellipses per camera pair as calculated from the covariance matrix – 3D view

6.5 Comparing the PMM to the CMM by Segmenting the Calibration Object

The use of back-projection of the calibration object to confirm system accuracy can be dangerous and should be used with due caution.

The problem is that it is essentially using the data that defined the system to test the system. Calibration finds the homography that linearizes the capture volume, based on some form of regression optimization between images points and corresponding world points. From the inherent nature of the linearization, regions nearer the calibration points behave more linearly than regions further from the calibration points. It is because of this behaviour that using the calibration points to test the system accuracy can give a false impression of the system accuracy.

Back-projection of the calibration object is a useful tool in that if the resulting error is large then there is definitely a problem with the system. If the resulting error is low, then proceed forward with caution.

So far the PMM has been compared to itself (through back-projection) and to the VICON. The shortfall with comparing the PMM to the VICON is that even though the VICON is of known accuracy, with a baseline being the CMM, the comparison does not indicate if the PMM is more or less accurate the VICON in terms of the baseline, merely that it is within a certain accuracy of the VICON.

Ultimately the PMM needs to be compared to the CMM directly, through world points that are not part of the calibration and that adequately cover the capture volume. Unfortunately a second calibration object could not be constructed before the completion of this thesis and this meant that a comparison of PMM to CMM would not be possible. Given that the preceding sections indicate that the PMM is of acceptable accuracy for the physiotherapy study, even if the worst case is assumed for the PMM-VICON comparison, not being able to compare the PMM to the CMM leaves certain doubt on the results and casts a blemish on the thesis. Worried about the impact of not having a comparison between PMM and CMM, the problem was discussed with Prof. Schreve, the supervisor. From the discussion it was realized that there is a way that the calibration object can be used to perform as both a calibration object and a comparison object without encountering the problem of bias.

6.5.1 Dividing the Calibration Object into Subsets

Having realized that the pyramid is a collection of points, it meant that the calibration object could be divided into two distinct subsets on the same object. Using one subset to calibrate and one subset as comparison, on a per camera and per camera pair basis.

Given the set of points in the calibration object as

$$S_{cal} = \{S_1,\, S_2 \colon s_i \in S_1,\, s_i \notin S_2\}, \tag{6.13}$$

with corresponding image points $IP_{cal} = \{IP_1,\, IP_2\}$, and CMM world points $WP_{cal} = \{X_1, X_2\}$.

The system can be calibrated using the subsets IP_1 and X_1 to find the homo-

graphies for the relevant cameras:

$$H_{cam_i} = f\left(IP_{1_i}, X_{1_i}\right) \tag{6.14}$$

Then the subset IP_2 can be reconstructed to get $\tilde{X}_2$:

$$IP_{2_i}, IP_{2_j} \xmapsto[H_{cam_j}]{H_{cam_i}} \tilde{X}_2 \tag{6.15}$$

The error of the system relative to the CMM can now be given as,

$$\varepsilon_{PMM_2} = \left\{\left\|\tilde{X}_2 - X_2\right\|\right\} \tag{6.16}$$

6.5.2 Results of Reconstructing the Subset

The images and selected images points of the calibration object are the same points used to calculated the back-projection results and the PMM-VICON comparison results. The data is without the E layer of the object and thus consists of all points $A - D$. Further the calibration subset consisted of all the even numbered points in each layer, while the comparison subset consisted of all the odd numbered points in each layer.

The system was calibrated with the even subset, the odd subset was reconstructed and compared with the corresponding CMM measurements for those points, the results are presented in Table 6.13:

Table 6.13: Statistical data comparing the PMM to CMM; calculated by reconstructing all odd numbered points of the pyramid calibration object

PMM	ε_x	ε_y	ε_z
$\boldsymbol{\mu_{Odd}}$	2.18 mm	0.61 mm	0.74 mm
$\boldsymbol{\sigma_{Odd}}$	1.69 mm	0.52 mm	0.43 mm
$\boldsymbol{SAE_{Odd}}$	43.61 mm	12.11 mm	14.87 mm
$\boldsymbol{\#_{Odd}}$	20	20	20

These values are close to those of the back-projection given in Table 6.9, for Y and Z axes, while in the X-axis the PMM-CMM comparison shows approximately 1 mm improvement in both mean and standard deviation over the values of the back-projection (approximately a 30% improvement). However the back-project

statistics are based on the entire calibration object of 40 points and this could be a reason for the larger difference in X-axis values.

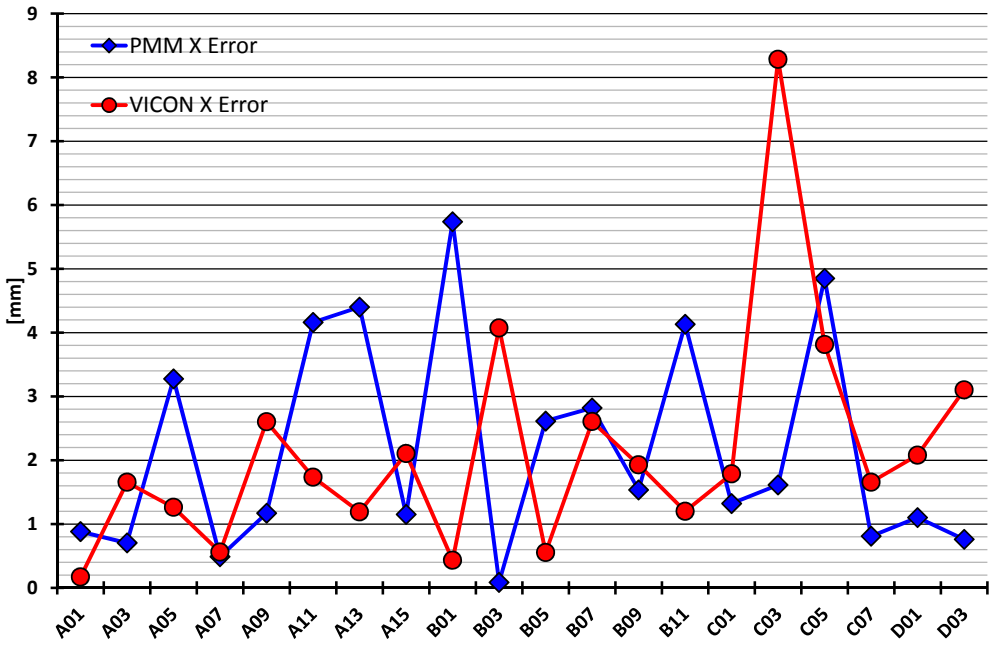
As the VICON was directly compared to the CMM using the pyramid calibration object, the data from the VICON capture was divided into odd and even subsets. The statistics for the odd subsect was calculated so that the PMM and VICON could be compared along side each other relative to the CMM. The statistics for the VICON-CMM odd subset is presented in Table 6.14:

Table 6.14: Statistical data comparing the PMM to CMM; calculated by reconstructing all odd numbered points of the pyramid calibration object

VICON	ε_x	ε_y	ε_z
μ_{Odd}	2.14 mm	2.09 mm	1.14 mm
σ_{Odd}	1.79 mm	1.42 mm	1.07 mm
SAE_{Odd}	42.79 mm	41.84 mm	22.72 mm
$\#_{Odd}$	20	20	20

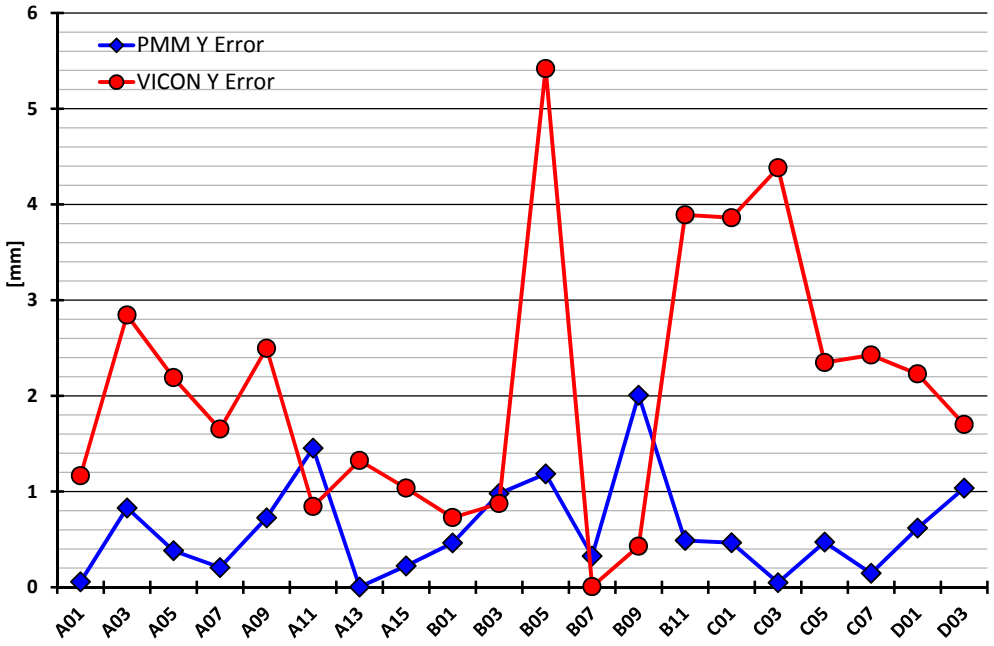
The PMM-CMM results are comparable in accuracy to the VICON-CMM results for the same subset. In the X-axis there is little difference between mean values, with the PMM is 1.9% worse than the VICON. However, in the Y-axis the PMM is 71% improved compared to the VICON and in the Z-axis the PMM is 34% improved over the VICON. The standard deviations present greater improvements of the PMM over VICON, with X-axis improving by 5.8%, Y-axis by 63.6% and Z-axis improving by 59.3%.

The error data from the PMM-CMM and VICON-CMM for the odd subset is plotted per axis in Figure 6.10, Figure 6.11 and Figure 6.12:



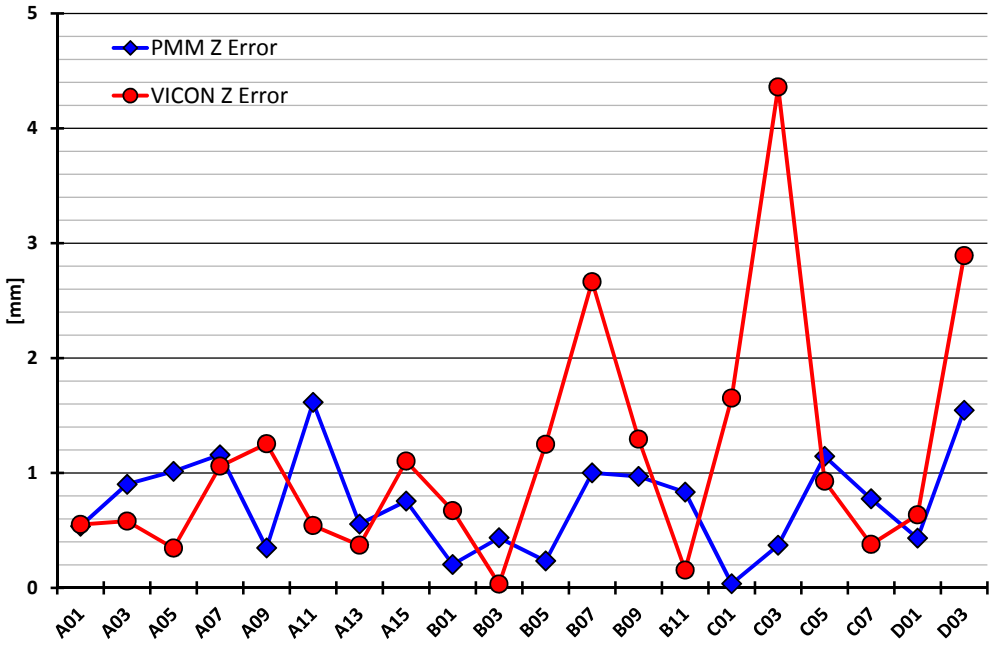
(a) Error along X-axis

Figure 6.10: X-axis error for the PMM and VICON vs. CMM (odd subset)



(a) Error along Y-axis

Figure 6.11: Y-axis error for the PMM and VICON vs. CMM (odd subset)



(a) Error along Z-axis

Figure 6.12: Z-axis error for the PMM and VICON vs. CMM (odd subset)

While this method does show appealing results, the method is not perfect.

- The calibration of 20 points over a volume 700 mm^3 does not represent the true accuracy of the object that usually has 40 points covering the same volume.
- Even though the two subsets are distinct, where the Comparison subset is not used to calculate the calibration of the system; due to the design of the pyramid calibration object the two subsets do sit in relatively close proximity to each other (50 mm to 100 mm) and do lie close to common planes. Ideally a comparison object should be dissimilar to the calibration object so that any accidental mathematical similarity between the two may be minimized.
- There are not many data points from which to calculate the statistics.

That said; the comparison subset does span the calibration volume evenly. Ultimately the two subsets are independent and thus the values do provide a direct comparison between the PMM and the CCM. An accuracy test with an independent object will still need to be performed, but this is certainly an improvement to the accuracy calculations than relying on a comparison to the VICON and a back-projection of the calibration object.

6.6 Concluding Remarks on the Experiment

For the setup of the PMM in the mannequin test and in the physiotherapy study, points were detected from 2 m to 3 m away. This data and these plots illustrate that the PMM, as it is currently designed and implemented in the lowest base form, is sufficiently accurate for the purposes of the physiotherapists.

In fact, the accuracy exceeds initially expectations. For the PMM to run in the school classroom environment, the cameras cannot be placed in the optimal positions. The cameras are close together, while far from the capture subject – leading to closer alignment of the cameras' focal axes, as evident from the large error in the x-axis. Due to these factors, it was initially thought that accuracy would be around 5 mm mean error with a 5 mm standard deviation (roughly half the marker diameter) when compared to the VICON — this was the preliminary goal to be achieved. The system exceeded this goal in the comparison with the VICON and showed that the calibration object methods for error estimation are valid and an improvement on the comparison with the VICON. This is further corroborated by using subsets of the calibration object to more accurately compare the PMM to the CMM, a test which gives better insight to the accuracy of the PMM. While all these are positive facts from the experiment, the PMM must still be validated against an totally independent object that tests the full range of the capture volume in more detail.

The data also illustrates that camera pairs that include C_1 exhibit more error than the other pairs. This, coupled with the pixel errors shown in Table 6.11, indicate that camera C_1 should be investigated as a source of increased uncertainty in the system.

The covariance error ellipses presented indicate the over-all accuracy of the system, and are a good visual representation of the system error. However using them to pin-point exactly which areas of the system require improvements is not straight-forward. While it can be seen that X-axis errors dominate the covariance error ellipses, it is only over cross-checking to camera placement that it can be attributed to the fact that most of the camera's optical axes lie closer to the X-axis than any other axis. This fact cannot be deciphered directly form the error ellipse. As a result a system with a large error ellipse could have one cause that is creating significant error, or multiple causes of small errors. This means that all possible factors need to be investigated and tested to determine how changes in those as-

pects effect the resulting error ellipse. To this end it is important to develop a list of *usual suspects*, aspects of the system that are usually suspected to be sources of significant error; aspects such as camera placement, distortion correction, poor calibration and so forth.

Chapter 7

Conclusion and Recommendations

7.1 Conclusions

A physical machine and supporting system, the Posture Measurement Machine (PMM), was successfully designed, implemented and field-tested to meet the original design specifications of the physiotherapists within the specified constraints.

The PMM has proven to be portable. Considering that it has been transported to (and around) twenty-six schools over a twelve-month period without incident or complications, its portability is evident.

The PMM has been proven to be accurate. With average axis errors ($\bar{\varepsilon}_{axis}$) of between 0.64 mm and 3.02 mm, these values are within the 5 mm tolerance set as an initial aim of the design. Furthermore, the standard deviations (σ_{axis}) for the axes are all below the 5 mm initial aim, with values between 0.6 mm and 2.65 mm.

The PMM has been proven to be a low-cost system. With a rough cost of R30 000.00 for all the cameras, the laptop, software and hardware, when compared to the currently available commercial vision measurement systems, costing between R100 000.00 (MotionStar) and R3.4 million (VICON).

Improvements to the system, such as the development of the new pyramid calibration object, allow for a 360° placement of cameras around a capture volume without the need to recalculate world-origin points due to the calibration object being moved.

Both lens distortion correction and sub-pixel accuracy enhancement were crucial improvements to the system, allowing the system to achieve the desired accuracy.

In the error model, the Jacobian of the angular function was analytically de-

rived and the whole model subsequently empirically proven, through numerical simulation, to be an accurate predictor of angular uncertainty to within 0.02° of arc. This makes the error model a useful design tool for assessing the use of the PMM as an angular measuring device for a given experiment.

The development of the system shows that the machine can serve a dual purpose as both a prototype tool capable of commercial development for use by physiotherapists, and as a research platform for further scientific research on the theory and application of computer vision.

All the aims of this Master's thesis were achieved successfully and the findings exceed expectations.

7.2 Shortcomings

While all the aims of the Master's thesis were achieved successfully, there still remained problems within the project that could not be dealt with adequately within the framework of this thesis. These include:

- The JPEG compression of images needs to be evaluated due to JPEG being a lossy algorithm, leading to the degradation of image quality. A lossless format such as PNG should be investigated as an alternative.

- Camera C_1 shows high pixel error, performing poorly in camera-pair triangulation. Further investigation of the images captured by C_1 revealed excessive noise in the images, as well as pronounced barrel distortion that could not be corrected successfully because of this noise. The replacement of this camera would improve the overall performance of the system.

- The software is not yet user-friendly, as there is no user interface. The creation of an interactive user interface would greatly benefit the system.

- The number of cameras in the system limits the viewing directions for capture. Currently there are only five cameras in the system that view the subject from the left and right sides. As a result, the views are unbalanced and it is difficult to guarantee that sufficient camera pairs will capture all markers on the subject.

7.3 Future Considerations

Further research should look at:

- The construction of the pyramid calibration object to be more rigid and robust, for instance through the use of hollow aluminium tubing to replace the wooden dowels.
- Additional cameras in the system to allow for greater flexibility in selecting viewpoints. The benefits for accuracy of increasing the number of cameras in the system should also be investigated.
- Cameras should be mounted independently on individual tripods, as opposed to multiple cameras on a crossbar. The benefits of this would be increased flexibility in the positioning of the cameras.
- The use of infra-red lighting and infra-red pass filters on the cameras. This would avoid the influence of variable ambient light conditions.
- The statistical nature of the uncertainties of point positions and distances in 3D.
- The feasibility of the commercialization of the system.

These are thought to be important points, the investigation of which would add value to the system and to the field of computer vision.

Appendices

Appendix A

Angle Definitions

A.1 3*D* Sitting Postural Angles as given by Physiotherapy

These are the angles as given by physiotherapist Yolandi Brink, based on [34; 35].

Head Flexion The angle between a line drawn form the *Cyclops* to the *OCI* and the *vertical axis*.

Neck Flexion The angle between a line drawn from the *OCI* to the *C7 SP* and the *vertical axis* in the *sagittal plane*.

Cranio-cervical angle The angle between a line drawn from the *Cyclops* to the *OCI* to the *C7 SP*.

Cervico-thoracic angle The angle between a line drawn from the *OCI* to the *C7 SP* to the *T5 SP*.

Trunk Flexion The angle between a line drawn from the *C7 SP* to the mid-point of the *greater trochanters* and the *vertical axis*.

Head Lateral Bending The lateral angle between a line drawn from the *OCI* to the *trachus of the ear* and the vertical line through the *OCI* (negative to the left).

Neck Lateral Bending The angle between a line from the *OCI* to the *C7 SP* and the vertical axis through *C7* in the sagittal plane.

Head Rotation The angle between a line drawn from the *OCI* to the *Cyclops* and the *anterior axis* in the transverse plane.

Thoracic Trunk Rotation The angle between a line drawn from the *sternum* to the the *T5 SP* and the *anterior axis* in the transverse plane (negative to the left).

A.2 Engineering Definitions

A.2.1 Mathematical Definitions of the Angles

These are the mathematical interpretations of the angles given in §A.1. The unit vectors $\left\{\hat{i}, \hat{j}, \hat{k}\right\}$ correspond to body-centred x-, y-, z-axes respectively.

1. $A_1 = \angle\left\{\hat{k} \mapsto \ell_1\right\}$
   - $\ell_1 = line : OCI \rightarrowtail Cyc$
   - $Cyc = \frac{LCanth+RCanth}{2}$
   - $OCI = \frac{LTrach+RTrach}{2}$
2. $A_2 = \angle\left\{\hat{k} \mapsto \ell_2\right\} : Sagittal\ plane$
   - $\ell_2 = line : C7 \rightarrowtail OCI$
3. $A_3 = \angle\{\ell_1 \mapsto -\ell_2\}$
4. $A_4 = \angle\{\ell_2 \mapsto -\ell_3\}$
   - $\ell_3 = line : C7 \rightarrowtail T5$
5. $A_5 = \angle\left\{\hat{k} \mapsto \ell_4\right\}$
   - $\ell_4 = line : MGT \rightarrowtail C7$
   - $MGT = \frac{LHip+RHip}{2}$
6. $A_6 = \angle\{\ell_5 \mapsto \ell_6\}$
   - $\ell_5 = line : LTrach \rightarrowtail RTrach$
   - $\ell_6 = line : \hat{k}$
7. $A_7 = \angle\{\ell_7 \mapsto \ell_2\} : Coronal\ plane$
   - $\ell_7 = line : \hat{k}$
8. $A_8 = \angle\{\ell_8 \mapsto \ell_1\} : Transverse\ plane$

- $\ell_8 = line : \hat{j}$

9. $A_9 = \angle\{\ell_{10} \mapsto \ell_9\} : Transverse\ plane$
   - $\ell_9 = line : T5 \rightarrowtail Strn$
   - $\ell_{10} = line : \hat{j}$

Appendix B

Reference Frame Transformations

Given two Cartesian coordinate reference systems, ***O*** and ***L***, it is desirable to have a transformation to convert points referenced in the one system to the other system.

For the rest of this explanation it will be assumed that all points are in ***O*** frame (unless otherwise stated), and that the desired transformation maps from ***L*** frame to ***O*** frame. Figure B.1 show two such generic systems. The vector $\vec{T}_r$ is the translation vector and is the position of the ***L*** frame origin within the ***O*** frame. The points $\vec{x}_1$, $\vec{x}_2$ and $\vec{x}_3$ are points in ***O*** that specifically form the desired unit vectors in ***L***.

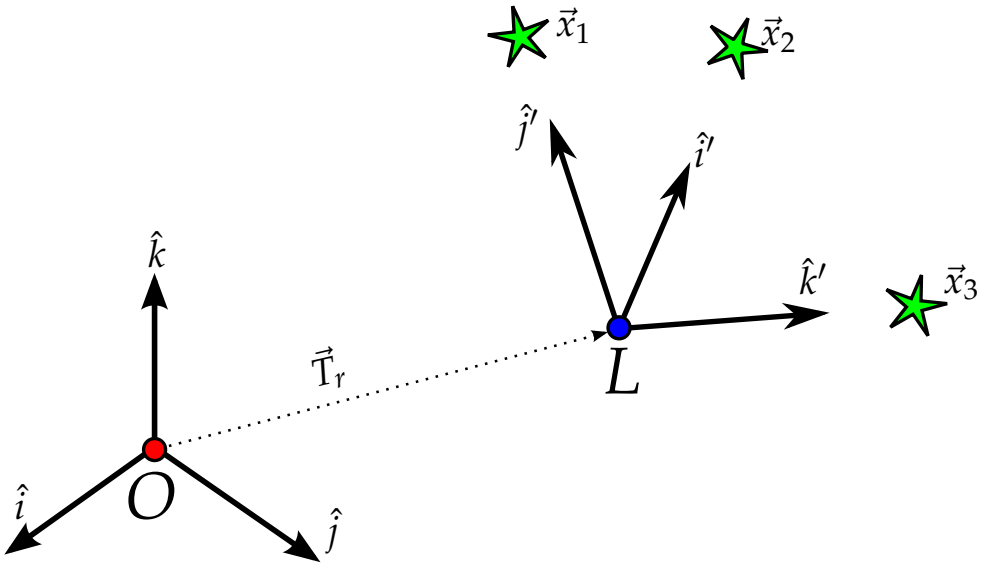


Figure B.1: Two Cartesian coordinate reference systems, ***O*** & ***L***

Thus the unit vectors of $\boldsymbol{L}$ in $\boldsymbol{O}$ are:

$$\begin{aligned} {}^{O}\hat{i}' &= \left\|\vec{x}_2 - \vec{T}_r\right\| \\ &\mapsto \langle \mathbf{1, 0, 0} \rangle = {}^{L}\hat{i}' \end{aligned} \tag{B.1}$$

$$\begin{aligned} {}^{O}\hat{j}' &= \left\|\vec{x}_1 - \vec{T}_r\right\| \\ &\mapsto \langle \mathbf{0, 1, 0} \rangle = {}^{L}\hat{j}' \end{aligned} \tag{B.2}$$

$$\begin{aligned} {}^{O}\hat{k}' &= \left\|\vec{x}_3 - \vec{T}_r\right\| \\ &\mapsto \langle \mathbf{0, 0, 1} \rangle = {}^{L}\hat{k}' \end{aligned} \tag{B.3}$$

Note that the translation vector was subtracted from the three points. The origins of the two coordinate systems *must* be at the same point, namely $\vec{\emptyset}$. This means that for all points in $\boldsymbol{L}$ that are to be transformed to $\boldsymbol{O}$, the origin offset must be removed through the use of the translation vector $\vec{T}_r$. Figure B.2 shows the translated frames.

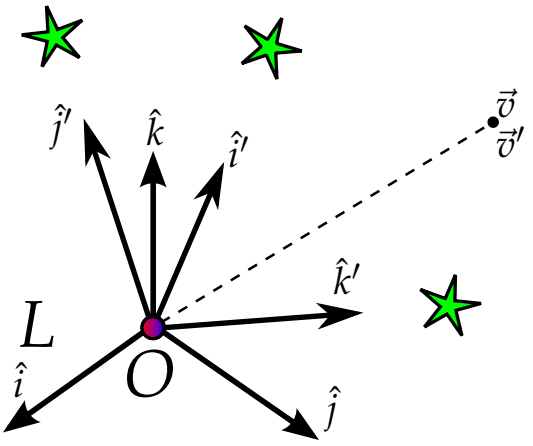


Figure B.2: The $\boldsymbol{L}$ system origin translated onto the $\boldsymbol{O}$ system origin

Once this origin offset is removed, the transformation becomes a rotational-scaling matrix that will act upon the points to be transformed. This can be derived with the use of the point vector, $\vec{v}$. As a vector in $\boldsymbol{O}$, with components,

$${}^{O}\vec{v} = v_x\hat{i} + v_y\hat{j} + v_z\hat{k}. \tag{B.4}$$

The components v_x, etc. are the perpendicular projection of $\vec{v}$ onto the corresponding unit vectors, such that:

$$v_x = \vec{v} \cdot \hat{i}, etc. \tag{B.5}$$

This means that ${}^{L}\vec{v}$ can be found by projecting ${}^{O}\vec{v}$ onto the unit vectors of $\boldsymbol{L}$, as in;

$$
\begin{aligned}
{}^{L}v_x &= {}^{O}\vec{v}\cdot\hat{i}' = \left(v_x\hat{i}+v_y\hat{j}+v_z\hat{k}\right)\cdot\hat{i}' = v_x\hat{i}\cdot\hat{i}'+v_y\hat{j}\cdot\hat{i}'+v_z\hat{k}\cdot\hat{i}' && \text{(B.6)}\\
{}^{L}v_y &= {}^{O}\vec{v}\cdot\hat{j}' = \left(v_x\hat{i}+v_y\hat{j}+v_z\hat{k}\right)\cdot\hat{j}' = v_x\hat{i}\cdot\hat{j}'+v_y\hat{j}\cdot\hat{j}'+v_z\hat{k}\cdot\hat{j}' && \text{(B.7)}\\
{}^{L}v_z &= {}^{O}\vec{v}\cdot\hat{i}' = \left(v_x\hat{i}+v_y\hat{j}+v_z\hat{k}\right)\cdot\hat{k}' = v_x\hat{i}\cdot\hat{k}'+v_y\hat{j}\cdot\hat{k}'+v_z\hat{k}\cdot\hat{k}' && \text{(B.8)}
\end{aligned}
$$

which can be written in matrix form as:

$$
\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}_L = \begin{bmatrix} \hat{i}\cdot\hat{i}' & \hat{j}\cdot\hat{i}' & \hat{k}\cdot\hat{i}' \\ \hat{i}\cdot\hat{j}' & \hat{j}\cdot\hat{j}' & \hat{k}\cdot\hat{j}' \\ \hat{i}\cdot\hat{k}' & \hat{j}\cdot\hat{k}' & \hat{k}\cdot\hat{k}' \end{bmatrix}\cdot\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}_O \qquad \text{(B.9)}
$$

$$
\Rightarrow {}^{L}\vec{v} = \boldsymbol{T}_{O/L}\cdot{}^{O}\vec{v} \qquad \text{(B.10)}
$$

Thus there is a matrix that transforms from $\boldsymbol{O}$ to $\boldsymbol{L}$.

The full transformation of a point in $\boldsymbol{L}$ to $\boldsymbol{O}$ is thus:

$$
{}^{L}\vec{X} = \boldsymbol{T}_{O/L}\cdot{}^{O}\left(\vec{X}-\vec{T}_r\right) \qquad \text{(B.11)}
$$

By inspection it can be seen that $\boldsymbol{T}_{O/L}$ is a symmetric matrix and therefore the reverse transformation matrix can be given by:

$$
\boldsymbol{T}_{L/O} = \left(\boldsymbol{T}_{O/L}\right)^{-1} = \left(\boldsymbol{T}_{O/L}\right)^{\mathsf{T}} \qquad \text{(B.12)}
$$

The reverse transformation of a point from $\boldsymbol{L}$ to $\boldsymbol{O}$ is given by:

$$
{}^{O}\vec{X} = \boldsymbol{T}_{L/O}\cdot{}^{L}\vec{X}+{}^{O}\vec{T}_r \qquad \text{(B.13)}
$$

Appendix C

Pyramid Calibration Object Ground Truth World Points

Here follows a table of pyramid object ground truth world points as measured on the computer measuring machine (CMM) by Prof. Schreve.

The x-axis runs from ***A*04** to ***A*12**. The y-axis is perpendicular to the x-axis and runs through ***A*00**. The z-axis is the cross product of the x- and y-axes. The origin is the intersection of the three axes. The pyramid object is measured relative to that definition.

Table C.1: Pyramid object world points as captured by the CMM (all values are in millimetres)

NM#	X	Y	Z	NM#	X	Y	Z
***A*00**	0	327.929	0	***C*00**	-1.383	167.908	290.254
***A*01**	-81.381	248.49	-0.061	***C*01**	-82.654	84.306	289.669
***A*02**	-164.507	167.028	0.146	***C*02**	-162.421	0.485	290.574
***A*03**	-247.592	84.961	-0.196	***C*03**	-80.159	-81.291	288.453
***A*04**	-328.274	0	0	***C*04**	-1.964	-165.04	290.326
***A*05**	-247.755	-81.922	0.199	***C*05**	81.521	-83.91	289.191
***A*06**	-165.357	-164.325	-1.292	***C*06**	166.634	1.506	289.299
***A*07**	-82.189	-247.199	-0.671	***C*07**	83.433	85.459	289.916
***A*08**	-2.145	-331.998	-2.154	***D*00**	-0.483	85.721	452.854
***A*09**	81.208	-246.574	0.069	***D*01**	-80.436	2.057	450.253
***A*10**	164.874	-165.214	-1.137	***D*02**	-3.955	-83.248	449.057
***A*11**	247.968	-82.746	-1.208	***D*03**	82.443	-0.125	450.061
***A*12**	330.205	0	0	***E*00**	-3.441	9.531	609.021
***A*13**	249.829	82.991	-1.009				
***A*14**	166.958	165.01	-1.497				
***A*15**	83.637	248.344	-0.687				
***B*00**	0.862	248.064	130.104				
***B*01**	-80.753	166.649	129.046				
***B*02**	-166.246	84.151	131.162				
***B*03**	-247.051	0.509	131.113				
***B*04**	-166.002	-82.004	129.045				
***B*05**	-83.294	-167.189	128.045				
***B*06**	0.878	-251.578	130.387				
***B*07**	84.845	-164.274	132.385				
***B*08**	164.245	-83.16	128.928				
***B*09**	246.815	-0.514	128.841				
***B*10**	164.882	81.048	126.761				
***B*11**	82.782	165.618	131.154				

List of References

[1] Prins, Y.: *The aetiology of upper quadrant musculoskeletal pain in high school learners using desktop computers: a prospective study*. Thesis (MScPhysio (Physiotherapy)), University of Stellenbosch, March 2008.

[2] Hartley, R.I. and Zisserman, A.: *Multiple View Geometry in Computer Vision*. 2nd edn. Cambridge University Press, 2004.

[3] Faugeras, O.: *Three-Dimensional Computer Vision: A Geometric Viewpoint*. The MIT Press, 1993.

[4] Whittle, M.: Calibration and performance of a 3-dimensional television system for kinematic analysis. *Journal of Biomechanics*, vol. 15, no. 3, pp. 185–196, 1982.

[5] Weng, J., Huang, T. and Ahuja, N.: Motion and structure from two perspective views: Algorithms, error analysis, and error estimation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, no. 5, pp. 451–476, 1989.

[6] Yang, S., Bhanu, B. and Mourikis, A.: Error model for scene reconstruction from motion and stereo. In: *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. IEEE, 2010.

[7] Chapron, M., Cocquerez, J. and Liu, N.: Precision of camera calibration and stereovision performed by standard cameras and image digitizer. In: *Pattern Recognition, 1992. Vol. I. Conference A: Computer Vision and Applications, Proceedings., 11th IAPR International Conference on*, pp. 704–707. IEEE, 2002.

[8] Pedersini, F., Sarti, A. and Tubaro, S.: Accurate and simple geometric calibration of multi-camera systems. *Signal Processing*, vol. 77, no. 3, pp. 309–334, 1999.

[9] Song, L., Wang, M., Lu, L. and Jing Huan, H.: High precision camera calibration in vision measurement. *Optics & Laser Technology*, vol. 39, no. 7, pp. 1413–1420, 2007.

[10] Zhou, J., Hao, Y., Zhu, F. and He, L.: The Effect of Camera Calibration Space on Visual Pose's Precision. In: *Natural Computation, 2007. ICNC 2007. Third International Conference on*, vol. 2, pp. 194–198. IEEE, 2007.

[11] Zhou, G., Tang, X. and Yuan, B.: Distortion model selecting and accuracy evaluation for CCD camera calibration. In: *Signal Processing, 1996., 3rd International Conference on*, vol. 2, pp. 875–878. IEEE, 1996.

[12] Hartley, R.: Self-calibration of Stationary Cameras. *International Journal of Computer Vision*, vol. 22, no. 1, pp. 5–23, February 1997.

[13] Shashua, A. and Navab, N.: Relative affine structure: Theory and application to 3D reconstruction from perspective views. In: *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*, pp. 483–489. IEEE, 2002.

[14] Wu, F., Hu, Z. and Zhu, H.: Camera calibration with moving one-dimensional objects. *Pattern Recognition*, vol. 38, no. 5, pp. 755–765, 2005.

[15] Trobina, M.: Error Model of a Coded-light Range sensor Communication Technology Laboratory. Tech. Rep., Image Science Group, ETH-Zentrum, 1995.

[16] Cheong, L. and Peh, C.: Depth distortion under calibration uncertainty. *Computer Vision and Image Understanding*, vol. 93, no. 3, pp. 221–244, 2004.

[17] Weng, J., Cohen, P. and Herniou, M.: Camera calibration with distortion models and accuracy evaluation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 14, no. 10, pp. 965–980, 1992.

[18] Mayoral, R., Lera, G. and Pérez-Ilzarbe, M.: Evaluation of correspondence errors for stereo. *Image and Vision Computing*, vol. 24, no. 12, pp. 1288–1300, 2006.

[19] Chowdhury, A. and Chellappa, R.: Statistical error propagation in 3D modeling from monocular video. In: *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW'03. Conference on*, vol. 8, p. 89. IEEE, 2008.

[20] Haralick, R.: Propagating covariance in computer vision. In: *Pattern Recognition, Proceedings of the 12th IAPR International Conference on Computer Vision & Image Processing*, vol. 1, pp. 493–498. IEEE, 1994.

[21] Seitz, S., Curless, B., Diebel, J., Scharstein, D. and Szeliski, R.: A comparison and evaluation of multi-view stereo reconstruction algorithms. In: *Computer Vision and*

Pattern Recognition, 2006 IEEE Computer Society Conference on, vol. 1, pp. 519–528. IEEE, 2006.

[22] Fusiello, A., Trucco, E. and Verri, A.: A compact algorithm for rectification of stereo pairs. *Machine Vision and Applications*, vol. 12, no. 1, pp. 16–22, 2000.

[23] Yuan, X. and Ryd, L.: Accuracy analysis for RSA: a computer simulation study on 3D marker reconstruction. *Journal of Biomechanics*, vol. 33, no. 4, pp. 493–498, 2000.

[24] Olague, G. and Mohr, R.: Optimal camera placement for accurate reconstruction. *Pattern Recognition*, vol. 35, no. 4, pp. 927–944, 2002.

[25] Schreve, K.: pyMultiCam User Guide, 2009. Dept. Mechanical and Mechatronic Engineering, Stellenbosch University.

[26] Article 22: Maximum number of IEEE-1394 cameras on a single 1394 bus. Tech. Rep., Point Grey Research, 2008.

[27] Herbst, B. and Hunter, K.: Computer Vision: Notes for TW892, 2009. Dept. Applied Mathematics, Stellenbosch University.

[28] Chen, L., Armstrong, C. and Raftopoulos, D.: An investigation on the accuracy of three-dimensional space reconstruction using the direct linear transformation technique. *Journal of Biomechanics*, vol. 27, no. 4, pp. 493–500, 1994.

[29] Gonzalez, R.C. and Woods, R.E.: *Digital Image Processing*. 3rd edn. Prentice-Hall, Inc., 2006.

[30] Arras, K.O.: An introduction to error propagation: Derivation, meaning and examples of $\boldsymbol{C_y} = \boldsymbol{F_x}\boldsymbol{C_x}\boldsymbol{F_x}^{\mathsf{T}}$. Technical Report EPFL-ASL-TR-98-01 R3, Swiss Federal Institute of Technology, Lausanne, 1998.

[31] Brink, Y.: Risk Factors Predictive of Neck, Shoulder & Arm Pain amid High School Students using Desktop Computers, 2009–2011. South African Medical Research Council Ethics nr: N08/08/209.

[32] Portney, L. and Watkins, M.: *Foundations of Clinical Research: Applications to Practice*. Pearson/Prentice Hall, 2009.

[33] Montgomery, D.C. and Runger, G.C.: *Applied Statistics and Probability for Engineers*. 4th edn. Wiley & Sons, 2007.

[34] Straker, L., Burgess-Limerick, R., Pollock, C., Murray, K., Netto, K., Coleman, J. and Skoss, R.: The impact of computer display height and desk design on 3D posture during information technology work by young adults. *Journal of Electromyography and Kinesiology*, vol. 18, no. 2, pp. 336–349, 2008.

[35] Straker, L., Maslen, B., Burgess-Limerick, R. and Pollock, C.: Children have less variable postures and muscle activities when using new electronic information technology compared with old paper-based information technology. *Journal of Electromyography and Kinesiology*, vol. 19, no. 2, pp. e132–e143, 2009.

[36] Bradski, G. and Kaehler, A.: *Learning OpenCV*. O'Reilly Media, Inc., 2008.

[37] Csurka, G., Zeller, C., Zhang, Z. and Faugeras, O.: Characterizing the uncertainty of the fundamental matrix. *Computer Vision and Image Understanding*, vol. 68, no. 1, pp. 18–36, 1997.

[38] Pinker, S.: *How the Mind Works*. Penguin Books, 1997.

[39] Poole, D.: *Linear Algebra: A Modern Introduction*. 2nd edn. Thomson Brooks/Cole, 2006.

[40] Remondino, F. and Fraser, C.: Digital camera calibration methods: considerations and comparisons. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 36, no. 5, pp. 266–272, 2006.

[41] Schreve, K.: Analytical Method for Estimating the Covariance Matrix. Tech. Rep., Dept. Mechanical and Mechatronic Engineering, Stellenbosch University, 2009.

[42] Stegmann, M.B.: *Image Warping*. Technical University of Denmark, October 2001.

[43] Zhang, Z.: Determining the epipolar geometry and its uncertainty: A review. *International Journal of Computer Vision*, vol. 27, no. 2, pp. 161–195, 1998.

[44] Zucchelli, M. and Kosecka, J.: Motion bias and structure distortion induced by calibration errors. In: *Proceedings of British Machine Vision Conference (BMVC), Manchester, UK*, pp. 663–672. Citeseer, 2001.